Screening of bamboo species available in West Bengal, India to assess their utility as lignocellulosic biomass resources

Submission of 5000 Words Synopsis

Submitted by: Subhadeep Biswas

Ph. D Registration No. – R-18RD18120142

Enrollment No. – PU/PHD/DLS/2018/010

Supervisor: Dr. Malay Das

Plant Genomics Laboratory, Department of Life Sciences,

Presidency University,

Kolkata.

Abstract

Bamboos, a category of non-timber grasses, have immense potential to be used in ligno-cellulose (LCB) based biofuel industry primarily due to their high growth rate, which leads to rapid accumulation of LCBs, high fibre contents, requirement of less agronomic input, existence of a wide genetic pool, and having no conflict with food crops. Despite this, there haven't been sufficient investigations to study the anatomy and biochemistry of bamboo's cellulose and lignin deposition. There remains a serious dearth of understanding, in particular, on the variations of cellulose and lignin levels that exist at the germplasm level.

Therefore, in this study, seven abundantly growing and widely distributed bamboo species of West Bengal, India, namely *Bambusa balcooa, B. tulda, B. bambos, B. nutans, B. striata, Dendrocalamus giganteus,* and *D. strictus* were investigated to assess their potential as LCB donors. In order to determine the chemical composition of LCBs obtained from the internode tissues of these species, this study utilized Attenuated Total Reflection-Fourier Transform Infrared (ATR-FTIR) Spectroscopy, Thermal Gravimetric Analysis (TGA), biochemical assessments and anatomical features. Findings of this study indicated that intensity of stain specific for cellulose and lignin could be a useful, yet cheap marker to screen a big collection of bamboo germplasms and to identify superior genotype. *B. bambos* was identified as the reference bamboo for ligno-cellulose research in future. Also, in this study three *Cellulose Synthesis A* (*CesA*) genes regulating secondary cell wall formation had been identified and sequence characterized from the tropical bamboo *B. tulda*.

Table of Contents

1. Introduction……1-2

2. Objectives (as specified in approved 1000 word)……..……..2-3

3. Methodology……3-13

3.1. Field survey and selection of bamboo species……3-4

3.2. Taxonomic identification of the selected bamboo species…….4-5

3.3. Collection of tissues for anatomical studies..……..5

3.4. Study of vascular anatomy by tissue sectioning, staining and subsequent microscopic observations……..…….……5

3.5. Selection of anatomical parameters and ImageJ analysis……..……..5

3.6. Collection of internode tissues for analytical and biochemical evaluations…….5-6

3.7. Fourier transform infrared spectroscopy (FTIR)……..…6

3.8. Estimation of moisture and α-cellulose contents of internode tissues……..…...6-7

3.9. Estimation of acid insoluble lignin contents in internode tissues……..……7

3.10. Thermogravimetric analysis (TGA)……..……...7

3.11. Cellulose crystallinity via XRD……..……...7

3.12. Statistical analysis……8

3.13. Collection of Plant Tissues for isolation of Nucleic Acids……..……..8

3.14. Retrieving gene and amino acid sequences from different databases for bioinformatics comparisons……..……8

3.15. Isolation of Genomic DNA……..……..8-9

3.16. Isolation of total RNA……..……..9

3.17. Preparation of cDNA libraries……..……9

3.18. Primer designing for gene amplifications and expression analyses…..……..9-11

3.19. PCR amplification of targeted genes from *B. tulda*……11

3.20. Gel purification of amplified PCR products……..……11

3.21. Preparation of Calcium competent *E. coli* Dh5α cells……11

3.22. Cloning of the amplified PCR products and transformation……..……12

3.23. Confirmation of successful transformed colonies via Colony PCR……..…12

3.24. Isolation of plasmids, Sanger's sequencing and data analyses……...13

3.25. Prediction of Gene model structures……..12-13

3.26. Analyses on the phylogenetic relationships of the identified genes with sequences from related monocot genomes……..13

3.27. Gene expression analyses using real time qRT-PCR method…..……...13

3.28. *In silico* identification of motifs……..……13

4. Results……..13-32

Objective 1: Comparisons of vascular anatomy of commonly growing bamboo species/populations of West Bengal, India to predict differences in cellulose and lignin contents

4.1. Variations in cellulose and lignin contents across bamboo species/genotypes assessed by the four anatomical parameters……..……14-18

Objective 2: Biochemical estimation of α-cellulose and acid insoluble lignin contents among selected bamboo species/populations

4.2. Chemical composition of grass LCBs were revealed by the FT-IR analysis...18-21

4.3. Assessing thermal stability to identify bamboo species having promising biofuel potential……..……...21-22

4.4. Measuring cellulose crystallinity index for the purpose of cell wall characterization……..……23

4.5. Biochemical levels of α-cellulose and acid insoluble lignin (AIL) contents vary among bamboo species/genotypes……..23-25

4.6. Statistical analysis combining anatomical and biochemical data…..……25-27

Objective 3: Identification, phylogenetic comparisons and expression characterization of selected *CesA* homologs in bamboo

4.7. *In silico* identification of *CesA* family genes from all the sequenced genomes of monocotyledonous plants……..……...27

4.8. Selection of candidate *CesA* genes involved in the secondary cell wall formation……27-28

4.9. Identification and molecular characterization of three SCW *CesA* gene homolog from *B. tulda*……..28-29

4.10. Study on the phylogenetic relationship of *B. tulda CesA* gene in comparison to the homologous sequences from related monocot species……..29-30

4.11. Comparative analyses of predicted gene model structures in relation to those of similar monocotyledonous species……30-31

4.12. *In silico* identification identification of domains present in CesA protein homologs of closely related monocot species……..31-32

5. Discussion…..……..32-34

6. Significance and Future direction……34

6. References……..……....34-38

1. Introduction:

Due to the global energy crisis, increasing industrial development and rapid population growth, search for renewable and sustainable energy based on non-food, lignocellulosic biomass (LCB) is an international priority. Lignocellulosic biomass is primarily composed of three polymers, which are cellulose, hemicellulose and lignin. In addition, smaller amounts of other components like acetyl groups, minerals and phenolic substituents are often present (Isikgor and Becer 2015). Lignocellulose, the principal component of the plant cell walls, is mainly composed of cellulose (40-60% of the total dry weight), hemicellulose (20-40%), and lignin (10-25%, (Kang et al. 2014). Depending on the type of lignocellulosic biomass, these polymers are organized into complex non-uniform three-dimensional structures. Lignocellulose has a nature to resist degradation and this feature of lignocellulose primarily depends on the crystallinity of cellulose and hydrophobicity of lignin. In addition, hemicellulose also deters in releasing cellulose from the cellulose-hemicellulose-lignin complex (Agbor et al. 2011; Morais et al. 2012; Barakat et al. 2013). Usually, cellulose is surrounded by lignin. In terms of chemical structure, it is a β-glucan linear polymer of d-glucose, which are linked by the β-1, 4-glycosidic bonds. Due to its crystalline nature, usually cellulose structure is difficult to break without enzymatic hydrolysis (Ruel et al. 2012). The linear cellulosic chain is made of approximately 500-14000 D-glucose units. Approximately, 36 hydrogen-bonded glucan chains form insoluble microfibrils in the secondary cell wall (Zhao et al. 2012).

Lignocellulosic biomass (LCB), which are non-food in nature are used as feedstocks for second generation biofuels. Advantages of using non-food derived LCBs are many such as they are abundantly available, inexpensive and are available from a wide variety of sources. Most of the LCBs generally comes from agricultural wastes e.g., bagasse (sugarcane), cob (corn, sorghum, millet), husk (wheat, rice, sorghum, millet, coconut, coffee), straw (wheat, rice, sorghum, millet), hull (rice, coffee, nuts), stalk (rice, cotton), stover (corn, sorghum, millet), leaves (sorghum, millet), fronds (coconut), shells (coconut, peanut, nuts), forest biomasses obtained from the softwoods (conifers and gymnosperm trees), hardwoods (angiosperm trees), forest wastes such as sawdust, wood chips, pruning residues, slashes, branches from dead trees, forest thinning, non timber plants (bamboo), municipal solid wastes (MSW) and annual or perennial grasses (Zabed et al., 2016). Grasses such as switchgrass, Miscanthus, canary grass, giant reed, alfalfa, and *Pennisetum purpureum* are also being explored lately as potential energy plants (Yousuf et al., 2020).

Bamboo is a very useful plant species with a wide range of genetic variation and economic importance (Basak et al., 2021). The plant family is widely dispersed in Asia, America, and Africa and has over 1670 species belonging to 125 genera (Biswas et al., 2022). Bamboo has a tremendous potential to serve as a bioresource for LCB biomasses because of a variety of

factors, including their extremely quick clonal development, which causes the accumulation of substantial LCB biomasses in a little period of time (He et al., 2013), presence of excessive fibre (Das et al., 2005), lack of need for arable land, prolonged harvest season, and presence of a sizable but underutilized genetic pool (Das et al., 2008, Basak et al., 2021). Despite this potential, very little studies have been conducted to determine the amount and composition of cellulose-lignin contents in various bamboo species and genotypes.

*Cellulose synthase*s (*CesA*s), which occur in hexameric complexes in a rosette structure are located in the plasma membrane (Arioli et al. 1998), (Kimura et al. 1999). Studies on the reference plant *Arabidopsis thaliana* had identified 10 *CesA* genes, out of which functions of six genes have been identified by mutant analyses (Richmond 2000). In *Arabidopsis*, usually three distinct *CesA* genes are involved in primary cell wall development (*AtCesA1*, *AtCesA3* and *AtCesA6*), while the other three are associated with secondary cell wall development (*AtCesA4*, *AtCesA7*and *AtCesA8*; Maleki et al., 2020). A total of 19 *CesA* and 38 *Cellulose Synthase Like* (*Csl*) genes, which are related to cellulose biosynthesis machinery had been identified (Peng et al. 2013) in the temperate bamboo *Phyllostachys edulis*. Furthermore, molecular characterization of *CesA* and *Sucrose Synthase* (*SUSY)* genes have been carried out in *P*. *edulis* (Chen et al. 2010; Huang et al. 2020; Huang and Cheng 2019; Chiu et al. 2006).

Biosynthesis of cellulose is a complex process and is completed in a three steps (1) plasma membrane-associated *Sucrose Synthase* (*SUSY*) directly channelize the UDP-glucose substrate to cellulose synthesizing machinery (2) *CesA* polymerize glucose monomers into glucan chains while recycling liberated UDP back to *SUSY* and (3) Korrigan (*KOR*), which is a membrane associated cellulase, acts as an editor or monitor to convert glucan chain to cellulose microfibril (Delmer and Haigler 2002); (Mølhøj et al. 2002). In this way, mutual interaction among CesA, SUSY and KOR proteins appears to regulate cellulose biosynthesis in plants (Read and Bacic 2002).

The aim of this study was to investigate on bamboo species that are commonly growing in West Bengal, India by conducting anatomical, biochemical and physicochemical analysis to assess their potential for lignocellulosic biomass production in order to select a superior bamboo species for the biofuel industry on the basis of the above mentioned parameters. Additionally, three *CesA* genes related to secondary cell wall formation had also been characterized.

2. Objectives (as specified in approved 1000 word):

Objective 1. Comparisons of vascular anatomy of commonly growing bamboo species/populations of West Bengal, India to predict differences in cellulose and lignin contents

Objective 2. Biochemical estimation of α-cellulose and acid insoluble lignin contents among selected bamboo species/populations

Objective 3. Identification, phylogenetic comparisons and expression characterization of selected *CesA* homologs in bamboo

3. Research methodologies:

3.1. Field survey and selection of bamboo species

A thorough field investigation was conducted in different places of West Bengal, India to get an idea about the bamboo species diversity and vegetation density. Natural bamboo growing habitats located in Barasat, Shyamnagar, Madhyamgram, Newtown of North 24 Paraganas district, Baruipur of South 24 Paraganas district, Bandel of Hooghly district, Bagdogra, Rangia, Dudhia, Barakothi of Darjeeling district have been extensively explored to ultimately select seven bamboo species for this study. These were *Bambusa tulda* Roxb., *B. balcooa* Roxb., *B. bambos* (L.) Voss, *B. nutans* Wall. ex Munro*, B. striata* Lodd. ex Lindl., *Dendrocalamus giganteus* Munro and *Dendrocalamus strictus* (Roxb.) Nees. These seven bamboo species were located at the different bio-geographical regions **(Table 1)** having substantial divergence with respect to the altitudes, rainfall and climate conditions. Internode tissue samples were collected from physiologically mature bamboo plants having comparable growth related characteristics, such as the height obtained by the culms, the number of lateral branches emerging from the node, the colour of the culms, the presence or absence of culm sheaths, and the presence or absence of bristles in the culm sheath ring.

Table 1. Geographical locations from where samples were obtained for this study.

Bamboo species	Genotype	Place of sample collection	Latitude (°N)	Longitude (°E)	Altitude (Above sea level, m)
Bambusa balcooa	B.b.ban1	Bandel, Hooghly	22.93	88.35	16
	B.b.ban2	Bandel, Hooghly	22.94	88.29	16
	B.b.shm1	Shyamnagar, North 24 Paraganas	22.84	88.40	02
B. tulda	B.t.ban1	Bandel, Hooghly	22.93	88.35	16
	B.t.ban2	Bandel, Hooghly	22.91	88.35	16
	B.t.shm1	Shyamnagar, North 24 Paraganas	22.83	88.41	02
B. bambos	B.bs.ban1	Bandel, Hooghly	22.93	88.35	16
	B.bs.ban2	Bandel, Hooghly	22.92	88.35	16

	B.bs.ban3	Bandel, Hooghly	22.92	88.35	16
B. nutans	B.n.nb1	Barakothi, Darjeeling	26.48	88.14	1226
	B.n.nb2	Panighata, Darjeeling	26.72	88.34	1226
	B.n.nb3	Panighata, Darjeeling	26.80	88.23	1226
B. striata	B.s.eco1	Ecopark, Newtown, North 24 Paraganas	22.60	88.46	01
	B.s.eco2	Ecopark, Newtown, North 24 Paraganas	22.60	88.46	01
	B.s.eco3	Ecopark, Newtown, North 24 Paraganas	22.61	88.46	01
Dendrocalamus giganteus	D.g.nb1	Bagdogra, Darjeeling	26.71	88.33	136
	D.g.nb2	Bagdogra, Darjeeling	26.71	88.34	136
	D.g.nb3	Rangia, Darjeeling	26.73	88.33	150
D. strictus	D.s.mdh1	Madhyamgram, North 24 Paraganas	22.68	88.46	15
	D.s.mdh2	Madhyamgram, North 24 Paraganas	22.68	88.46	15
	D.s.mdh3	Madhyamgram, North 24 Paraganas	22.68	88.46	15

3.2. Taxonomic identification of the selected bamboo species

The floral characteristics of angiosperms are usually used for taxonomic identification. Yet in bamboos, flowering is a highly irregular biological phenomenon and that even varies from species to species. For instance, *B. balcooa* flowers after 32-34 years, *B. tulda* after 48 years, and *B. bambos* after 30–50 years (Zheng et al., 2020). Due to such unavailability of floral features, these species were identified on the basis of 15 culm and 17 culm sheath characters, which have been reported before (Das et al., 2007). The fifteen culm characters used were plant height, diameter, length of the 5th internode, internode bending, culm colour, presence or absence of swollen node, nodal ring, nodal sheath scar, hairs at nodal ring, whether branches come out by piercing culm sheath or not, presence or absence of curved nodal branches, leaf size, modifications and striations. Similarly, seventeen culm sheath characters such as ratio of total length to breath at base, ciliate margin**,** presence or absence of pubescent adaxial and abaxial side, hair colour, number of hairs, ratio of total length vs blade length, shape of blade, blade reflexed or not, presence or absence of hairy blade margin, ligule margin, hairs on ligule, auricle, auricle continuous with blade or not, bristles, whether auricles fringed or not and sheath size (Das et al., 2008). All the voucher specimens of these seven bamboo species were submitted to the Calcutta University National

Herbarium (CUNH), which is affiliated to Botanical Survey of India (BSI), Shibpur (Accession numbers: 20047 – 20086, dated 25.11.2019).

3.3. Collection of tissues for anatomical studies

For the purpose of tissue-sectioning, various internodal tissues were harvested from 1st lateral branches of the above mentioned identified genotypes of all the seven bamboo species. The first lateral branch of the main culm, which was picked randomly for each genotype, served as the representative sample for that particular bamboo species. Each branch was subsequently separated into three sections i.e., apex, middle (mid), and basal (base) region. Each species was given consideration for three distinct genotypes. This led to a total of 63 tissue samples being used for the anatomical studies (7 species X 3 genotypes X 3 internode positions).

3.4. Study of vascular anatomy by tissue sectioning, staining and subsequent microscopic observations

An internode was divided into apex, middle and base and were stored in 10% Formaldehyde: 50% Alcohol: 5% Acetic Acid solution (FAA) fixative solution. Transverse, thin sectioning of each internodal tissue was done and stained with Fast green FCF (SRL, India) and Phloroglucinol-HCL (SRL, India). Fast green is a stain specific for cellulose (Vazquez-Cooz et al., 2002), while Phloroglucinol is specific for lignified tissues (Lin et al., 2002; Pradhan Mitra et al., 2008). Pictures were captured using a brightfield Nikon (Nikon, Japan) motorized inverted microscope Ti2-E, equipped with DS-Fi3 camera and a Ti2-D-LHLED LED lamp housing for Dia Illumination. Photos were taken using NIS-ELEMENTS AR Imaging software. Images were taken for each sample at magnifications of 4x, 10x, and 20x. A total of 378 pictures (63 samples x 2 stains x 3 magnifications = 378) were obtained and subjected to additional analysis.

3.5. Selection of anatomical parameters and ImageJ analysis

The anatomical parameters used for this study were a) total numbers of vascular bundles (VB) recorded at 4X magnification, b) lumen diameter of individual metaxylem recorded at 10X magnification, c) intensity of Fast Green stain (FG) and d) intensity of Phloroglucinol stains (PG). These anatomical data were analysed using the ImageJ software (Rueden et al. 2017). Stain intensity was measured using ImageJ and expressed as integrated density (ID) of that specified stain.

3.6. Collection of internode tissues for analytical and biochemical evaluations

For analytical and biochemical studies, internode tissues from the 1st lateral branches were harvested and collected from previously identified populations of all seven bamboo species. the total sample size was 63 (7 species X 3 genotypes X 3 internode positions). In order to

obtain dry biomass, freshly collected internode tissues were brought to the lab and placed in the 60° C dry oven for 48–72 hours, or until their weight is constant. They were then crushed in a mixer grinder to produce powdered sample, which was then packaged properly with labels and kept in a desiccator to prevent moisture absorption.

3.7. Fourier transform infrared spectroscopy (FTIR)

By detecting functional groups that are present in the examined samples, attenuated total reflectance - Fourier transform infrared spectroscopy (ATR-FTIR) provides information on the chemical composition of polymer materials. It is a non-destructive method that does not depend on chemical sample preparation and enables direct observation of samples in the solid state. It is superior than conventional technologies in several ways, including detection speed, sensitivity, precision, and higher spectrum resolution. Internode tissues obtained from 21 samples (7 species x 3 genotypes/species) were subjected to this analysis. Approximately 10 mg powder was used for each run. Using an FTIR spectrometer available at the Indian Institute of Engineering Science and Technology (IIEST), Shibpur, Howrah, unique infrared spectra were generated for each of these samples (Model no. FT IR - 4700LE, JASCO, Japan). Data were obtained from 24 running scans with a resolution of 4 cm^{-1}, which were used to obtain FTIR spectra in the mid-infrared light region, which covered the wavelength range of 4000 to 400 cm^{-1}. Data were analysed and spectrum graphs were produced Using OriginPro® 2022.

3.8. Estimation of moisture and α-cellulose contents of internode tissues

An entire internode was divided into apex, middle and base part for all the seven bamboo species. Moisture contents were estimated by considering the difference between fresh and dry biomass of harvested tissues. Fresh moisture content was estimated using the following equation

$$\text{Moisture Content \%} = \frac{(FW-DW)}{DW} * 100 \text{ [FW = Fresh Weight, DW = Dry Weight]}$$

Total α-cellulose contents of all bamboo species were estimated using the anthrone assay following standard procedures (Updegraff 1969; Bhattacharya et al. 2010). Around 500 mg of powdered internode tissues were utilized in this method. Then they were mixed to 3 ml of nitric acid: acetic acid (Merck, India) solution (1:8 v/v). Next, the sample was refluxed in hot water for 30 minutes. Colour of the sample was changed from dark brown to orangish yellow. After that, repeated washings were done with distilled water, followed by vortexing and centrifugation at 6000 rpm for 15 minutes. Each washing step eliminated lignin, hemicellulose and other possible contaminants. Supernatants were discarded after each wash and this

procedure was repeated till the colour of the discarded supernatant solution turned waxy white, the pellet was then dissolved in 5 ml of 67% sulfuric acid (v/v) (Merck, India). It was then diluted 100 times by combining it with pre-cooled 67% (v/v) sulphuric acid. From this diluted sample, 1 ml. was taken and afterwards combined with 3 ml. of chilled anthrone (0.2% in conc. H2SO4; SRL, India) reagent. The solution was heated for 10 minutes in boiling water, quickly chilled for 2 minutes on ice, and then incubated at room temperature for 10 minutes. Using a spectrophotometer (Shimadzu, Japan) light absorption was measured at 620 nm. The α-cellulose content was estimated using a standard curve of the cellulose solution (SRL, India) and was eventually expressed as the gram percentage of the dry weight.

3.9. Estimation of acid insoluble lignin contents in internode tissues

Total lignin is composed of acid insoluble lignin (AIL) and acid soluble lignin (ASL). However, out of these two forms, AIL contribute approximately 90-95% of the total lignin. In the pretreatment step, ASL can be excluded through acid treatment (e.g., 72% H2SO4) (Akinyemi et al., 2020), but not AIL and hence creates hindrance during saccharification. Therefore, Klason lignin or acid-insoluble lignin (AIL) content was determined in this study based on the patented protocol of the American Society for Testing and Materials [ASTM E1721 - 01(2015)].

3.10. Thermogravimetric analysis (TGA)

Thermogravimetric analysis (TGA) is used to evaluate the thermal stability of composing materials present in a sample**.** To examine the pattern of thermal degradation, about 10mg of oven-dried internode powder was put into the Discovery 650 SDT thermal analyser (TA instruments, USA). Samples were heated under nitrogen flow at a rate of 50 ml/min from ambient temperature to 900 °C. Heating rate was kept at 5° C/min. A mass spectrometer available at the Indian Institutes of Science Education and Research (IISER), Kolkata was used to continuously monitor the weight loss. The 1st. order derivative of all the TGA curves were used to obtain the Differential Thermo Gravimetry curves (DTG). TGA thermogram was obtained using the OriginPro® 2022 by analysing the raw data.

3.11. Cellulose crystallinity via XRD

The crystallinity indexes were estimated using x-ray diffraction (XRD) analysis. The analysis was performed with a Rigaku-Ultima III X-ray diffractometer with Cu Kα radiation source (λ= 1.5418 Å) operating at 40kV, 60mA. Samples were scanned from 2θ = 10° - 40° at a speed of 2°/min. The XRD machine available at the UGC-DAE CSR, Kolkata was used for this experiment. The crystallinity indexes of the biomasses were determined using Segal's method (Segal et al., 1959) as illustrated in following equation:

$$\text{Crystallinity Index (CrI) \%} = \frac{(I_{002}-I_{am})}{I_{002}} * 100$$

Where I_{002} is the intensity of main 002 peak at 2θ = 22.4° and I_{am} is the scattered intensity of amorphous portion at 2θ = 18°. Raw data were analysed and the spectra graphs were generated using the OriginPro® 2022.

3.12. Statistical analysis

Statistical analyses were done in order to compare among the anatomical and biochemical data obtained from the seven bamboo species. Pairwise correlation coefficient and Spearman's Rho statistical analysis were used to assess statistical level of correlation between the parameters. Based on all the 63 samples, a regression line was plotted based on the correlation analysis of biochemical estimation of cellulose and lignin concentration. Statistical correlation was also measured between the biochemically estimated cellulose and lignin content along with the four anatomical parameters. Further, by considering biochemical estimates of both cellulose and lignin along with the integrated density (ID) of FG and PG, scores were calculated both for the cellulose and lignin in each of the seven bamboo species.

3.13. Collection of Plant Tissues for isolation of Nucleic Acids

Vegetative tissues such as such as young leaf, internodes from emerging culm, middle aged developed culm, matured culm, leaf, culm sheath, root, rhizome and reproductive tissue such as late inflorescence buds were harvested from three populations of *B. tulda* in West Bengal, India, that were located in Shyamnagar (SHYM7, SHYM16, 22.38º N. 88.40º E) and Bandel (BNDL22, 22.93º N. 88.38º E), West Bengal, India.

3.14. Retrieving gene and amino acid sequences from different databases for bioinformatics comparisons

Different genomic databases such as PLAZA (https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v4_5_monocots), PHYTOZOME (https://phytozome.jgi.doe.gov/pz/portal.html), NCBI (https://www.ncbi.nlm.nih.gov) and Bamboo genome database (https://gendbe.computational.bio.uni-giessen.de/cgi-bin/login.cgi) were used to obtain gene, amino acid and upstream sequences for bioinformatics analysis such as Basic Local Alignment Search Tool (BLAST), pair wise alignment and multiple sequence alignment. *Arabidopsis thaliana* and *Oryza sativa* sequences were used to perform BLASTP analyses in order to obtain homologous sequences from the other related monocot species.

3.15. Isolation of Genomic DNA

Genomic DNAs were isolated from the young, folded leaves by using DNeasy Plant Mini Kit using manufacture's guidelines (Qiagen, Germany). Maximum 600 mg tissues were used for extraction. RNaseA treatment (100 mg/ml) was conducted and genomic DNAs were eluted in

50 µl DNase, RNase free water. Quality of isolated DNAs were assessed with a BioSpectrometer (Eppendorf, Germany) and 1.5% low EEO agarose (SRL, India) gel electrophoresis.

3.16. Isolation of total RNA

Isolation of total RNA from various tissue stages were performed by following the combination of TRIzol and RNeasy method (Das et al., 2016). Eluted RNA was treated with DNase I enzyme solution (Thermo Scientific, USA) to make it free of genomic DNA contamination. Quality and yield of the extracted total RNA was evaluated by the help of a biospectrometer (Eppendorf, Germany) and agarose formamide gel electrophoresis.

3.17. Preparation of cDNA libraries

The Verso cDNA synthesis kit (Thermo Fisher Scientific) was used for preparation of the cDNA libraries. To get the highest yield, 1 µg of total RNA was used for each library during the first strand cDNA synthesis. PCR cycle used for these reactions was 25 ºC for 10 mins, 42 ºC for 1hr and finally 70 ºC for 15 mins. Later, the amplified cDNAs were observed on agarose (SRL, India) gel electrophoresis.

3.18. Primer designing for gene amplifications and expression analyses

In order to amplify the selected genes from genomic DNA or cDNA, degenerate primers were designed based on the conserved regions of the homologous genes identified in related monocot species. Gene specific primers were designed from the amplified *B. tulda* sequences to perform gene expression analyses **(Table 2)**. Intron spanning or intron flanking primers were designed to avoid any possible amplification from contaminating genomic DNA. All the primers were designed using the Primer3 web tool (http://primer3.wi.mit.edu/). PCR amplified bands of desired length were excised from the agarose gel and eluted by using GeneJET Gel Extraction kit (Thermo Scientific, USA) following the manufacturer's guideline.

Table 2. Detailed information on primers designed for gene amplification and real time RT-qPCR analyses.

Gene	Sequence (5'-3')	Primer length (bp)	GC content (%)	Tm (ºC)	Amplicon length (bp)	Experimental purpose
CesA4	F - CTCTGCAAGGCCTGCCTCGAC	21	71.4	64.9	1518	PCR amplification of the target gene
	R - CGCATAATGATTACTGCCCTGTAAGGTG	28	50	61		
	F - ATGGAGTCGGCGGTCC	16	68.8	57.6	564	
	R - CTCCAGACACACTGCTAATGGTC	23	52.2	57.3		

	F - TTTGACAGATGCATATGAGCCACT	24	41.7	56.6	1783	
	R - TGTCAGAACAGCAGACACTCGAAC	24	50	58.9		
	F - GTTCGAGTGTCTGCTGTTCTG	21	52.4	55.8	1962	
	R - AGCACCATCGACTGCTGA	18	55.6	56.5		
	F - GCTATGTCCAATTCCCACAAAGA	23	43.5	55.3	152	Gene expression analyses by real time RT-qPCR
	R – TTGCCTATAGAAACAACAGCCA	22	40.9	54.4		
CesA7	F - ATGAAGAGGGAGTACGAGGAGTTCA	25	48	58.7	1704	PCR amplification of the target gene
	R - TTCACCGTGGGGATGATGA	19	52.6	56.1		
	F - ATGGACACCGGCTCGGTAA	19	57.9	58.8	1513	
	R - CACTTCATATGGACAATGCC	25	48	58.7		
	F - CCCTCGTTGCCTACTGTACCAT	22	54.5	58.6	740	
	R - TTAGCACTGGACTCCACACGG	21	57.1	59.1		
	F - GCAAGAAGGACAAGCTCG	18	55.6	53.8	190	Gene expression analyses by real time RT-qPCR
	R - GCGTGGAGGCGATGAACA	18	61.1	58.3		
CesA9	F - AGTGCAACACCCGCTACAAGCG	22	59.1	62.1	869	PCR amplification of the target gene
	R - TTCCTCCACAGCGGCTGCC	19	68.4	68.9		
	F - ATGGAGGCGAGCGCCG	16	75	61.8	347	
	R - TCGTCGTCGATGTTGAACTCGTG	23	52.2	59.1		
CESA9	F - CTACCCCGTGTCCGAGCC	18	72.2	60.5	2814	PCR amplification of the target gene
	R - GCATCTCCTTGTCACTGTCCATC	23	52.2	57.6		
	F - AGATGGTGACCTGCGACTGCTG	22	59.1	61.6	1507	
	R - TCAGCAGTTGATGCCACACTGC	22	54.5	60.4		
	F - TGCAGGTATGAGAGGGAAGG	20	55	56.2	656	
	R - TCACGGGAGACGTACACCA	19	57.9	58		
	F - AGATGGTGACCTGCGACTGCTG	22	59.1	61.6	256	
	R - GTCCACTTGAAGGCGTACAGCT	22	54.5	59.4		

	F - CTCGTGGAGAAGCAGAAGGTC	21	57.14	57.4	200	Gene expression analyses by real time RT-qPCR
	R - CTGCTGCGTCCTCTGCCCT	19	68.42	62.7		

3.19. PCR amplification of targeted genes from *B. tulda*

PCR amplifications were performed to amplify regions either from genomic DNA or cDNA of *B. tulda*. The amplifications were conducted by either using the Go Taq DNA Polymerase (Promega, USA) or high fidelity ExTaq (Takara, USA) DNA polymerase in the S1000 PCR detection system (Bio Rad, USA). The composition and concentrations of the PCR master mix was as follows: PCR Reaction buffer 10x, 1.5–2.5 mM MgCl2, 0.2 mM dNTPs, 0.25 mM forward and 0.25 mM reverse primer, 1 ng/µl genomic DNA, and 1.25 u/50 l ExTaq DNA polymerase (Takara, USA). The following cycle was used to carry out the PCR reactions: a 30-second initial denaturation at 98 °C, followed by 25–30 cycles of [98 °C for 10 seconds, 55–65 °C for 30 seconds, and 72 °C for 30–60 seconds per kb of PCR product, and a final extension at 72 °C for 3 minutes. Amplified PCR products were run on a 1.5% low EEO agarose gel electrophoresis (Lonza, USA) stained with 10mg/ml ethidium bromide (C21H20BrN3; SRL, India), and finally observed under UV illumination, and confirmed.

3.20. Gel purification of amplified PCR products

Thermo Scientific's GeneJET gel elution kit was used to elute the amplified PCR products from the agarose gel. Amplified PCR products were loaded in agarose gel stained with 10mg/ml ethidium bromide. By incubating in a water bath at 58 - 60 °C, band of desired molecular weight was dissolved into equal volume of binding buffer. After washing, it was passed through the GeneJET gel elution column. Finally, the product was eluted with 20µl nuclease free water.

3.21. Preparation of Calcium competent *E. coli* Dh5α cells

In order to perform cloning and transformation procedures, calcium competent *E. coli* Dh5α cells were prepared. In the beginning, *E. coli* Dh5α cells were grown overnight to be used as a primary culture and next day inoculated in 250 ml liquid LB medium, grown at 37 °C to use as secondary culture. The cell density was continuously monitored at 600 nm in a Spectrophotometer. The culture was maintained on ice once the density of cells reached 0.40. The bacterial cell pellet was collected by centrifugation at 8000 rpm for 5 min at 4 °C. This pellet was then dissolved in chilled 100 mM $MgCl_2$, and centrifuged once more for 5 minutes at 8000 rpm. Then, 100mM $CaCl_2$ solution was added to the $MgCl_2$ treated pellet, and they were centrifuged for 5 minutes at 4 °C at 8000 rpm. The solution was centrifuged once more after being re-suspended in 85mM CaCl2 in 15% glycerol. This final mixture of competent cells

was distributed equally in 2 ml. centrifuge tubes and were stored in a -80 °C refrigerator for future uses.

3.22. Cloning of the amplified PCR products and transformation

TA-based pGEM®-T Easy vector systems (Promega, USA) were used to clone the amplified PCR products having 3' A-tail. The process involved ligating PCR products: vector in a 3 (approx. 75ng.) :1 (approx. 25ng) ratio. The ligation mixture was incubated for overnight at 4 °C for TA cloning. Then, after adding it to the calcium-competent *E. coli* DH5α cells, the ligated products were maintained on ice for 30 minutes. The bacterial cells were subjected to heat shock at 42 °C for 45 seconds to induce transformation before being immediately transferred to ice. The transformed cells were then added with approximately 400 µL of liquid SOC medium for recovery, and they were then cultured at 37 °C at 130–140 rpm in a shaking incubator. These transformed cells were then plated on solid LB media with 1.6% agar and ampicillin (100 µg/ml; HiMedia, India), which is a selectable marker, was used for further selection. The X-Gal (5-Bromo-4-Chloro-3-Indolyl β-D-Galactopyranoside; 20 mg/ml, dissolved in formamide; SRL, India) and isopropyl - d-1-thiogalactopyranoside (IPTG; 0.1M; SRL, India) were also added to the solid medium for the purpose of selection. Plates were then kept at 37 °C overnight to encourage bacterial growth.

3.23. Confirmation of successful transformed colonies via Colony PCR

To identify putatively transformed colonies, approximately ten to twenty randomly chosen single colonies were added to separate PCR tubes containing PCR master mixture. From section 3.19 master mix's composition and the PCR cycle parameters were both followed. Procedure related to PCR amplification and visualization in agarose gel has already been discussed in the previous section.

3.24. Isolation of plasmids, Sanger's sequencing and data analyses

Putatively transformed colonies were cultured overnight at 37 °C in 10 ml of liquid LB medium having Ampicillin (100 µg/ml, HiMedia, India) as the selectable marker. Plasmid DNAs were isolated by GeneJET Plasmid Miniprep kit (Thermo Scientific, USA) using the manufacturer's guideline and finally eluted using 30µl of nuclease free water. Sequencing was done by standard Sanger's sequencing method. The plasmids were sequenced using primers that were both gene and vector specific. Trimmed nucleotide sequences were assembled by CAP3 software (https://doua.prabi.fr/ software/cap3). Finally, sequence data were uploaded to NCBI (http://www.ncbi.nlm.nih.gov/).

3.25. Prediction of Gene model structures

The online bioinformatics application tool Gene Structure Display Server (GSDS 2.0, http://gsds.gao-lab.org/) was used to build the gene models utilizing full length genomic and coding sequences (CDS).

3.26. Analyses on the phylogenetic relationships of the identified genes with sequences from related monocot genomes

To study the phylogenetic relationships of the selected gene homologs, full length deduced amino acid sequences of *B. tulda* were multiple aligned with homologous sequences from related monocot species using the MUSCLE programme (www.ebi.ac.uk/Tools/msa/muscle/) together. The Molecular Evolutionary Genetics Analysis (MEGA) version 11 programme was used to develop the Neighbour Joining (NJ) tree using its default settings (Kumar et al., 2016). To ensure that all the nodes had enough statistical support, bootstrap values of 1000 replicates were employed for all phylogenetic studies.

3.27. Gene expression analyses using real time qRT-PCR method

Expression analyses of the selected three *B. tulda CesA* genes in various tissues were studied using real time RT-qPCR analyses by using SsoAdvanced™ Universal SYBR® Green Supermix (BioRad, USA) in a CFX connect real-time PCR detection system. The final volume of the PCR mixture used was 10 µL, which contained 1X concentration of supermix buffer, 0.25 µM for each forward and reverse primers and 2 µL of cDNA (1/20th diluted from the stock). The PCR amplification cycles used were: 95 ºC for 30 s, followed by 40 cycles of 94 ºC for 10 s and at 55 ºC for 40 s. After completion of amplification, a dissociation curve (65–95 ºC, at increments of 0.5 ºC) analysis was performed to ensure absence of any primer-dimer. The relative changes in gene expression were calculated using the $2^{-\Delta\Delta Ct}$ method (Schmittgen and Livak 2008). The expression data were normalized using already identified *eIF4α* gene, which is a stable reference gene in bamboos (Chakraborty et al. 2019).

3.28. *In silico* identification of motifs

By using the Multiple EM for Motif Elicitation (MEME) version 5.3.0 tool, motifs found in the predicted proteins of *B. tulda* and other homologs were identified. Following parameters were used for this study: maximum number of motifs: 20, maximum width: ≤200 and minimum width: ≥6 and number of repetitions: any.

4. Results

Objective 1: Comparisons of vascular anatomy of commonly growing bamboo species/populations of West Bengal, India to predict differences in cellulose and lignin contents

4.1. Variations in cellulose and lignin contents across bamboo species/genotypes assessed by the four anatomical parameters

The first lateral branch of the main culm of each bamboo species was used for this study **(Fig. 1A** and **B)**. Transverse sectioning of the internode **(Fig. 1C)** of the seven bamboo species revealed their anatomical characters. The parenchymatous ground tissues surrounding the hollow central pith were evenly dispersed with collateral VBs. Diverse frequencies of VB were seen along different regions of the ground tissues as well as in different species. Usually, VB frequencies were higher around the culm perimeter than those in the pith cavity. Each vascular bundle unit is mainly made up of one to two protoxylems, two large metaxylem vessels, and a phloem, surrounded by the cellulose fiber cap **(Fig. 1D)**. Broadly, two types of vascular bundles were observed in the different species. Phloem and protoxylem elements of *B. balcooa*, *B. tulda* and *B. bambos* had sclerenchymatous fibre cap and termed as 'double broken-waist type' VB.

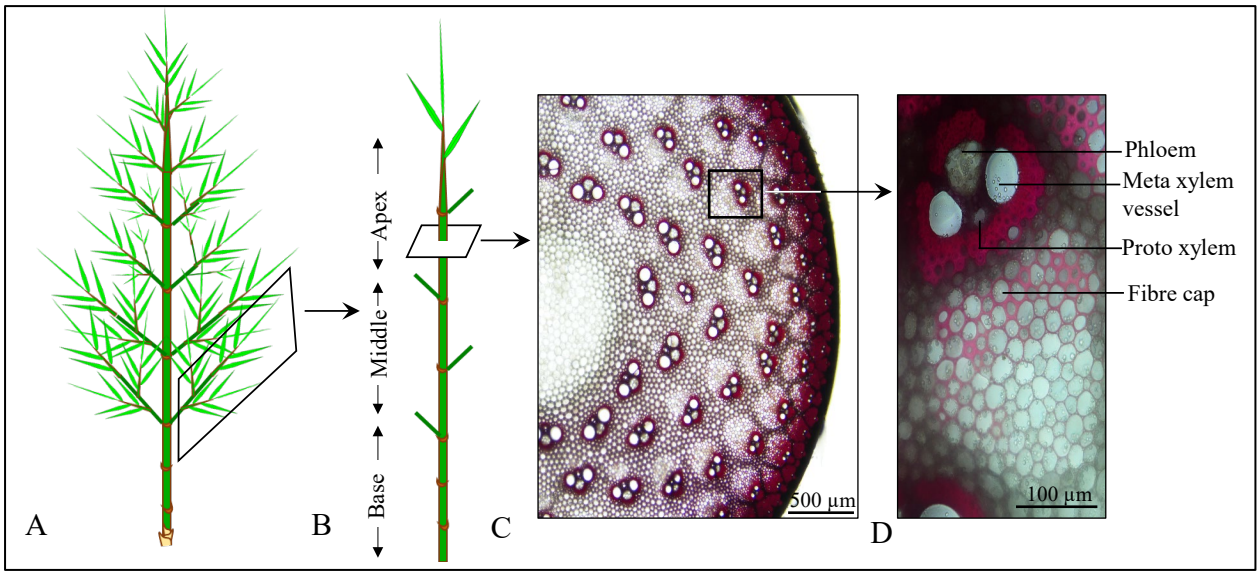

Fig. 1. Representative tissue stages and transverse section of internode tissues obtained from *B. tulda*. A. Schematic diagram of a bamboo culm **B.** Three regions of the lateral branches **C.** Transverse section of *B. tulda* internode. **D.** A single vascular bundle unit. The other four species only had fibre caps in their protoxylem and were referred as the "single broken-waist type" VBs. Due to significant deposition of lignocellulose biomasses, VBs observed around the culm periphery were fully mature and had either little or no metaxylem lumen. On the other hand, VBs, which were located near the pith cavity were much younger having larger lumen diameter. Four anatomical parameters were taken into consideration for this study to evaluate the usefulness of anatomical parameters for assessing cellulose and lignin contents in diverse bamboo species and genotypes. They were i. VB density or the number of VB per unit area; ii. metaxylem lumen diameter; iii. Integrated Density (ID) for FG;

and iv. Integrated Density (ID) for PG. Total 63 samples (7 species x 3 genotypes x 3 internode positions) were used to obtain data on all these anatomical characteristics. Data collected for all the anatomical parameters had been summarised in **Table 3**. Differences in VB numbers were identified among species. For instance, *B. nutans* (*B.n.nb2*) has the largest number of VBs, followed by *B. striata* (*B.st.eco3*) and *B. bambos* (*B.bs.ban1*).

Table 3. All the four anatomical parameters used to study the seven bamboo species.
Abbreviations used: ID = Integrated Density, FG= Fast Green, PG= Phloroglucinol.

Species	Genotypes	Total vascular bundle number	Metaxylem lumen diameter (µm)	ID of FG	ID of PG
Bambusa balcooa	B.b.ban1.Apex	102	47.72	3.80	7.45
	B.b.ban2.Apex	120	42.83	6.92	7.70
	B.b.shm1.Apex	164	44.11	5.04	9.47
	B.b.ban1.Mid	180	32.76	12.60	11.61
	B.b.ban2.Mid	124	40.50	9.28	12.94
	B.b.shm1.Mid	260	33.10	8.53	12.59
	B.b.ban1.Base	216	31.23	19.19	18.04
	B.b.ban2.Base	184	31.82	13.39	14.57
	B.b.shm1.Base	192	31.88	20.57	15.79
B. tulda	B.t.ban1.Apex	140	46.70	6.16	7.31
	B.t.ban2.Apex	124	51.85	4.57	6.57
	B.t.shm1.Apex	132	54.27	5.66	7.03

	B.t.ban1.Mid	156	36.77	9.77	12.45
	B.t.ban2.Mid	144	41.50	12.99	14.45
	B.t.shm1.Mid	156	42.70	15.16	11.49
	B.t.ban1.Base	220	30.39	14.98	17.89
	B.t.ban2.Base	208	31.07	20.51	20.20
	B.t.shm1.Base	200	30.51	22.61	17.11
B. bambos	B.bs.ban1.Apex	100	45.74	6.00	7.10
	B.bs.ban2.Apex	132	48.80	5.22	6.88
	B.bs.ban3.Apex	108	37.10	7.98	7.44
	B.bs.ban1.Mid	224	32.83	14.80	12.91
	B.bs.ban2.Mid	220	43.19	15.83	15.66
	B.bs.ban3.Mid	184	31.11	13.48	16.19
	B.bs.ban1.Base	260	24.34	20.31	14.19
	B.bs.ban2. Base	232	30.97	19.78	19.85
	B.bs.ban3.Base	192	23.70	20.58	20.72
B. nutans	B.n.nb1.Apex	170	33.17	9.27	8.85
	B.n.nb2.Apex	156	54.42	10.89	11.59
	B.n.nb3.Apex	208	56.43	11.50	14.09

	B.n.nb1.Mid	252	25.69	15.87	15.43
	B.n.nb2.Mid	220	40.91	12.49	14.50
	B.n.nb3.Mid	240	45.27	13.12	19.17
	B.n.nb1.Base	324	22.70	18.54	15.92
	B.n.nb2.Base	332	31.46	17.76	24.99
	B.n.nb3.Base	330	37.67	18.96	23.21
B. striata	B.st.eco1.Apex	213	56.32	5.16	6.17
	B.st.eco2.Apex	218	51.56	8.45	8.35
	B.st.eco3.Apex	198	56.15	7.30	6.80
	B.st.eco1.Mid	276	46.42	11.72	9.08
	B.st.eco2.Mid	234	41.31	10.49	11.20
	B.st.eco3.Mid	228	38.86	7.47	7.39
	B.st.eco1.Base	243	34.91	13.59	9.20
	B.st.eco2.Base	260	34.10	13.39	13.49
	B.st.eco3.Base	284	27.75	12.71	16.41
Dendrocalamus giganteus	D.g.nb1.Apex	144	54.39	12.85	17.01
	D.g.nb2.Apex	144	51.12	7.63	15.41
	D.g.nb3.Apex	164	54.40	12.70	18.39

	D.g.nb1.Mid	196	47.88	16.16	21.55
	D.g.nb2.Mid	164	52.66	14.21	14.93
	D.g.nb3.Mid	228	53.57	12.70	12.03
	D.g.nb1.Base	216	35.89	19.27	30.84
	D.g.nb2.Base	204	40.23	17.09	16.72
	D.g.nb3.Base	272	43.98	20.31	25.84
D. strictus	D.s.mdh1.Apex	110	54.96	5.53	9.80
	D.s.mdh2.Apex	138	58.31	8.02	9.22
	D.s.mdh3.Apex	122	55.68	6.40	11.57
	D.s.mdh1.Mid	201	34.51	11.18	10.76
	D.s.mdh2.Mid	237	40.59	9.29	9.34
	D.s.mdh3.Mid	162	35.73	11.65	13.76
	D.s.mdh1.Base	260	33.02	12.50	19.41
	D.s.mdh2.Base	276	25.85	10.12	11.03
	D.s.mdh3.Base	243	25.02	13.77	16.88

Objective 2: Biochemical estimation of α-cellulose and acid insoluble lignin contents among selected bamboo species/populations

4.2. Chemical composition of grass LCBs were revealed by the FT-IR analysis

The ATR-FTIR spectroscopy analysis was conducted to qualitatively asses the chemical composition of grass LCBs. The diagnostic peaks for the cellulose, hemicellulose and lignin

were obtained **(Table 4)** and these results were validated by comparing absorption peaks observed in other bamboos (Shi et al., 2019; Ling et al., 2021), oak (Kubovský et al., 2020) and other woody plants (Hosseinaei et al., 2012). Sharp peaks at around 3340 cm^{-1}, 2891 cm^{-1}, 1723 cm^{-1} **(Fig. 2 A-G)** and another sharp peak at 1030 cm^{-1} indicated dominance of cellulose and hemicellulose in the sample. The peaks obtained in the region of 1200 to 900 cm^{-1} were of cellulose and hemicellulose. These peaks include C-O(H/C) stretching from the alcohols/ether functionalities and 1-4 linked glycosidic bonds. Yet the prominent pointed top at 1156–59 cm^{-1} indicated that hemicellulose contains 1-3 linked xyloglucan molecules. Presence of lignin were also confirmed by the presence of aromatic lignin monomer peaks at ~1595-99 cm^{-1} and ~1420-24 cm^{-1}. The sharp peaks at ~897-99 cm^{-1} and ~830 cm^{-1} also revealed the presence of both the syringyl (S) and guaiacyl (G) monomer units.

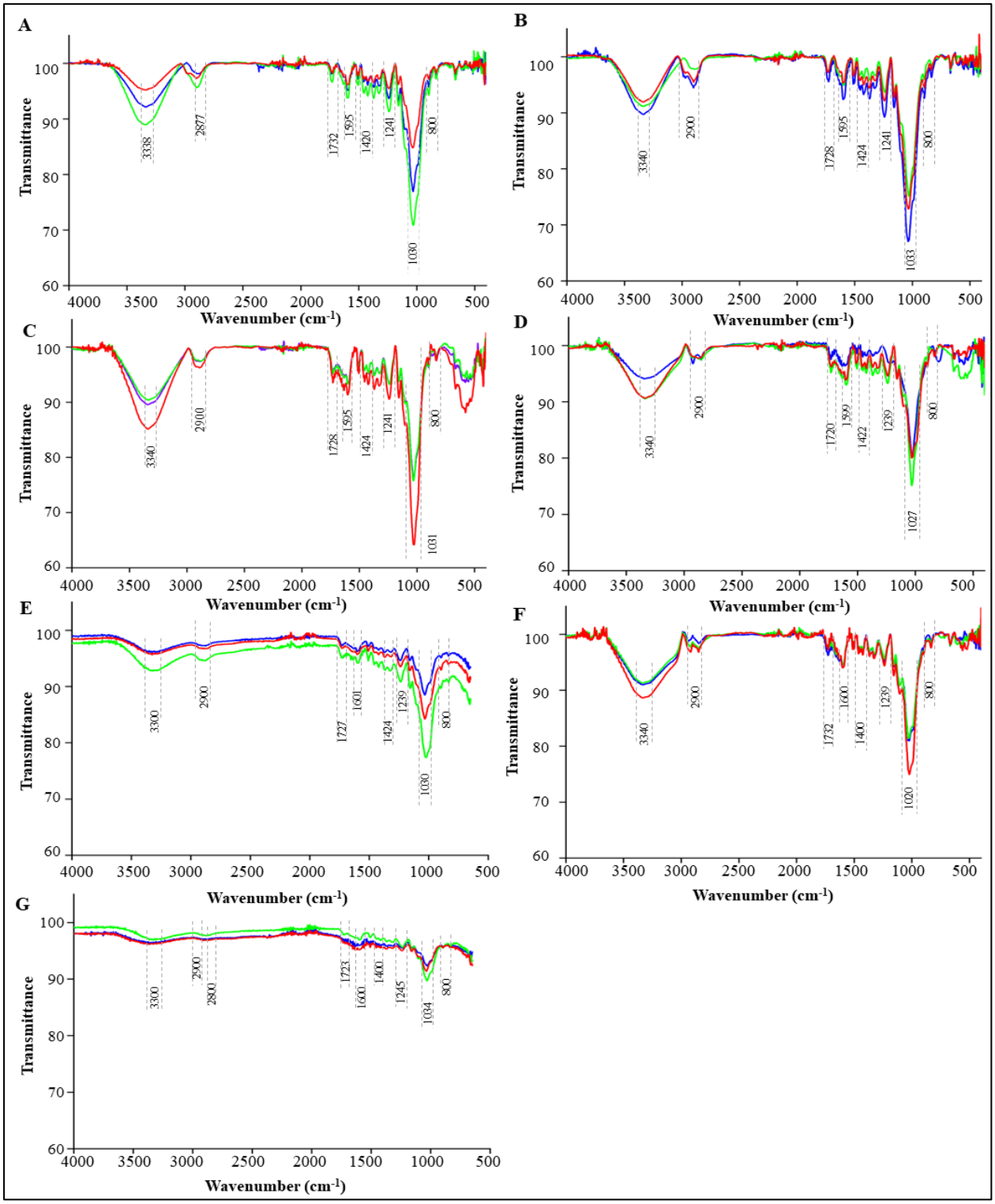

Fig. 2. ATR-FTIR spectra of seven bamboo species. A. *Bambusa nutans* B. *B. bambos* C. *B. balcooa* D. *B. tulda* E. *B. striata* F. *Dendrocalamus giganteus*, G. *D. strictus*. Transmittance is presented in Y-axis; wave number is presented in X-axis. For each species, data collected from 3 genoypes have been included and marked by distinct colours.

Table 4. Assignment of significant functional groups from the FTIR spectra.

Wavenumber (cm^{-1})	Vibration	Functionalities	Components
~3340 (strong and broad)	–O–H stretching	Presence of hydrogen bonded –OH	Cellulose, Hemicellulose, Lignin
~2923-92 and ~2891-98	$–CH_3$ and $–CH_2–$ asymmetric and symmetric stretching	Presence of open and ring moieties	Cellulose, Hemicellulose, Lignin
~2891-98	–(O=) C–H stretching	C–H stretching in aldehyde group	Cellulose, Hemicellulose, Lignin
~1723-32	–(H)C=O stretching	Aldehydic carbonyl groups	Cellulose, Hemicellulose, Lignin
~1027-33	C−O(H/C) stretching	in alcohols/ethers	Cellulose, Hemicellulose,
~1156-59	C−O–C stretching	in ethers of 1-3 linked Xyloglucan	Hemicellulose,
~1595-99 and ~1420-24	–(Ar) C=C stretching	Pair of stretching in Aromatic moieties	Lignin
~1239-41 and ~1027-33	C−O–C stretching	–(Ar) C–O–C two prominent peaks for Phenolic moieties	Lignin
~897-99 and ~830	Out of plane bending vibration of -C-H (Ar)	*ortho- and para-* substituted Aromatic moieties	Lignin

4.3. Assessing thermal stability to identify bamboo species having promising biofuel potential

Non oxidative thermal degradation analyzed by the TGA and DTG curves **(Fig. 3. A-B)** from internode tissues of the seven bamboo species revealed their thermal stability and demonstrated the three staged degradation of LCBs. In stage I, evaporation of moisture took place between 100-150 °C, which caused weight loss due to evaporation of moisture. Loss of weight ranged from 5.2% (*D. strictus*) to 8.6% (*B. bambos*). The majority of the pyrolysis of hemicellulose took place in stage II (250–300°C). This incurred loss of weight from 13.57% (*B. nutans*) to 21.45% (*B. bambos*). In stage III (300-350°C), primarily pyrolysis of cellulose occurred resulting

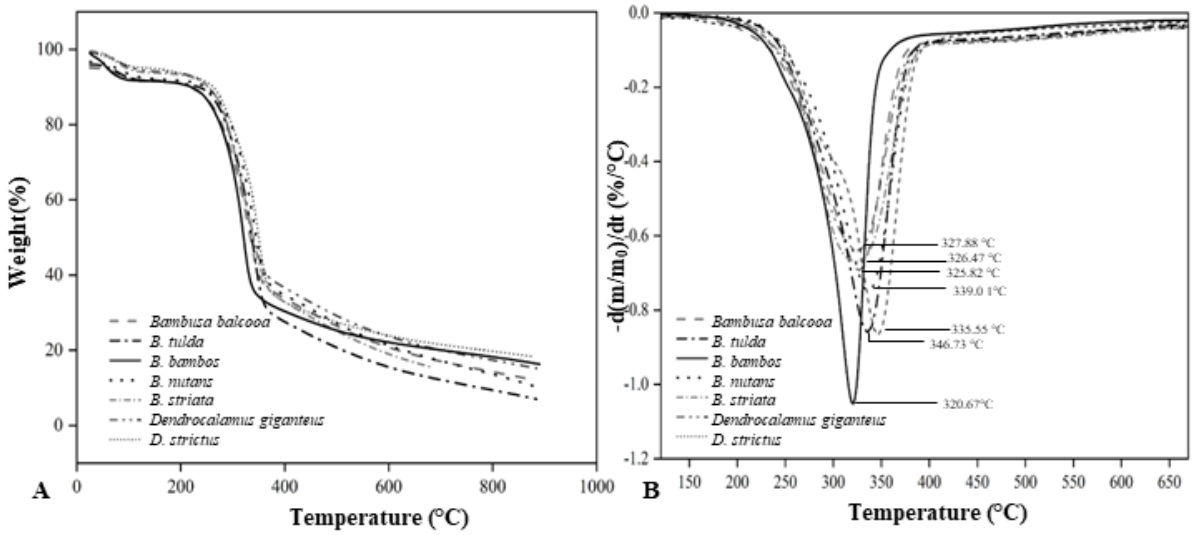

Fig. 3. A. Thermogram of seven bamboo species under nitrogen flow. B. DTG curve of seven bamboo species. The arrows indicate maximum decomposition temperature (T_{max}) for each species. Thermal stability was deduced by analysing the maximum decomposition temperature (T_{max}).

Table 5. TGA and XRD profile of seven bamboo species

Bamboo species	Moisture content (%)	Residual char at 900 °C (%)	Maximum decomposition temperature (T_{max})	Crytallinity index (CrI) (%)
Bambusa balcooa	8.3	5.01	327.88	41.46
B. tulda	8.59	4.75	335.55	41.12
B. bambos	8.6	4.12	320.67	41.05
B. nutans	7.82	7.13	339.01	44.56
B. striata	6.5	4.79	326.47	49.34
Dendrocalamus giganteus	5.92	4.76	325.82	49.29
D. strictus	5.2	3.01	346.73	50.21

Into the most significant weight loss, which ranged from 35.66% (*D. giganteus*) to 49.93% (*B. nutans*, **Table 5**) Degradation of lignin started occurring from approximately 400 °C until 800 °C. Finally, the remaining leftover material was known as residual char which was found lowest in *D. strictus* (3.01%) and highest in *B. nutans* (7.13%). *B. bambos* was the least stable species demonstrating maximum degradation at 320.67 °C and *D. strictus* was the most thermally stable bamboo demonstrating maximum degradation at 346.73 °C (**Table 5**, **Fig. 3.B**).

4.4. Measuring cellulose crystallinity index for the purpose of cell wall characterization

The efficacy of the cellulolytic enzymes to perform saccharification depends on cellulose crystallinity, which has a significant impact on the recalcitrance of cell walls. The recalcitrance increases with cellulose crystallinity index (CrI). The CrI was calculated via X-ray diffraction (XRD) method (**Fig. 4**) and following the Segal's method, the crystallinity index of seven bamboo species had been calculated. Maximum CrI was obtained in *D. strictus* **(50.21%; Table 5)** and *B. bambos* was the species having the least CrI **(41.05%, Table 5)**. Crystallinity index pattern of the bamboo species found to be comparable with decomposition temperature (T_{max}) patterns obtained from the TGA and DTG.

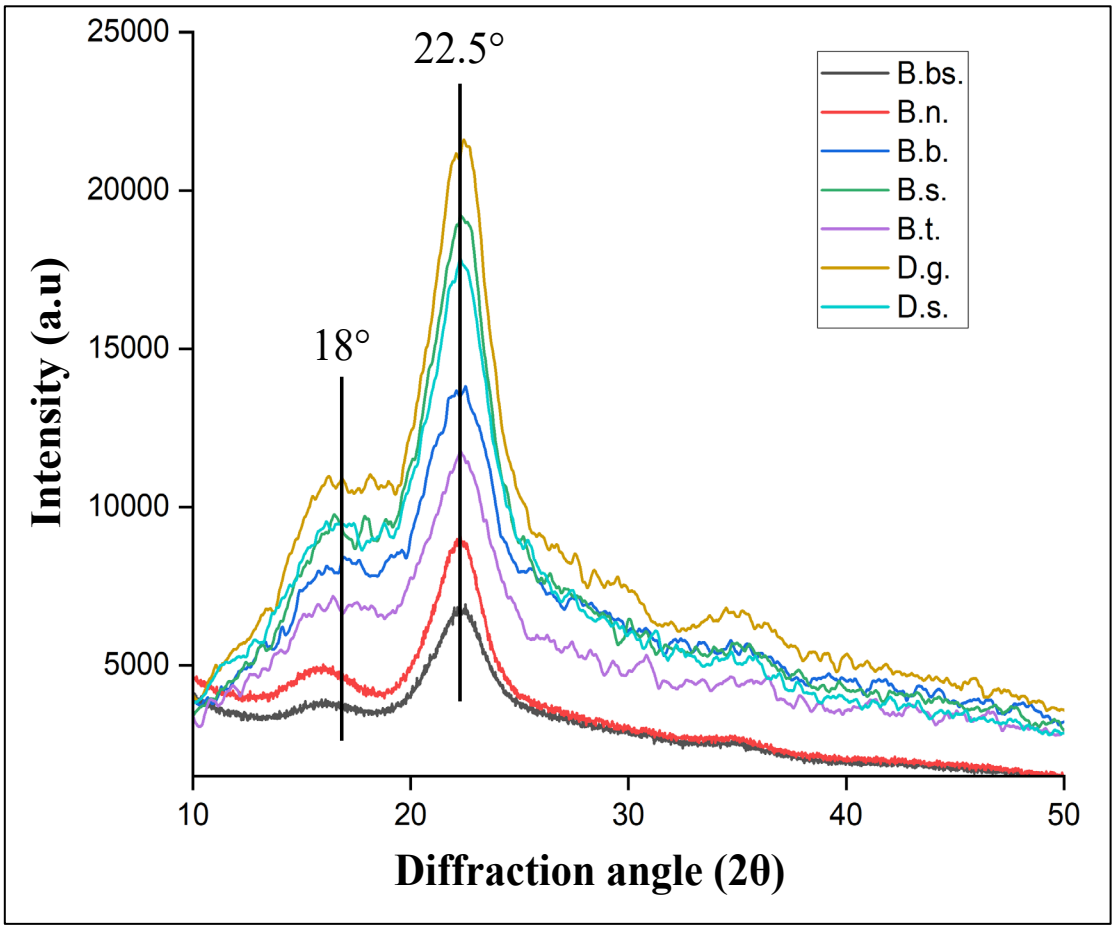

Fig. 4. X-ray Diffraction patterns of seven bamboo species showing peaks for amorphous and crystalline cellulose. Abbreviations used: B.bs. = *Bambusa bambos*, B.n. = *B. nutans*, B.b. = *B. balcooa*, B.s. = *B. striata*, B.t. = *B. tulda*, D.g. = *Dendrocalamus giganteus*, D.s. = *D. strictus*.

4.5. Biochemical levels of α-cellulose and acid insoluble lignin (AIL) contents vary among bamboo species/genotypes

Biochemical estimation of α-cellulose in the selected bamboos demonstrated variation both across species as well as genotypes **(Table 6)**. As per the biochemical estimations, *B. bambos* (*B.bs.ban3*) had the highest amount of cellulose at the basal internode (74.41 gm%), which was followed by *B. nutans* (*B.n.nb3*, 71.64 gm%) and *B. bambos* (*B.bs.ban2*, 65.36 gm%). In contrast, the two genotypes of *D. strictus* had the lowest levels of cellulose (*D.s.mdh3*: 25.08 gm% and *D.s.mdh1*: 26.82 gm%, **Table 6**), followed by *B. striata* (B.s.eco1, 27.55 gm%). When lignin content was assessed, *B. striata* had the lowest amount (16.12 gm%; *B.s.eco1*), followed by *D. strictus* (*D.s.mdh3,* 18.13 gm%; **Table 6**) and *D. giganteus* (*D.g.nb1,* 18.88 gm%; **Table 6**).However, , all the three *B. nutans* genotypes showed the highest levels of lignin in the basal internode (*B.n.nb2,* 25.30 gm%, *B.n.nb1,* 23.55 gm%, and *B.n.nb3,* 23.39 gm%).

Table 6. Biochemical estimation of the α-cellulose and acid insoluble lignin contents in seven bamboo species and three internode positions.

Bamboo species	Genotypes	α-cellulose content (gm%)			Acid insoluble lignin content (gm%)		
		Apical internode	Middle internode	Basal internode	Apical internode	Middle internode	Basal internode
Bambusa balcooa	*B.b.ban1*	23.69	33.38	57.27	13.52	17.37	20.84
B. balcooa	*B.b.ban2*	18.24	26.43	36.94	16.30	18.53	19.31
B. balcooa	*B.b.shm1*	16.56	31.78	51.98	15.22	16.63	21.40
B. tulda	*B.t.ban1*	19.18	23.69	62.89	17.19	20.07	21.64
B. tulda	*B.t.ban2*	14.59	25.43	39.01	16.32	20.33	21.15
B. tulda	*B.t.shm1*	17.46	37.07	60.86	17.31	19.11	22.42
B. bambos	*B.bs.ban1*	26.05	31.65	53.82	11.60	18.77	20.82
B. bambos	*B.bs.ban2*	19.45	31.79	65.36	16.94	20.81	21.21
B. bambos	*B.bs.ban3*	16.0	45.13	74.41	16.81	21.32	21.04
B. nutans	*B.n.nb1*	22.83	40.74	43.92	20.19	21.66	23.55
B. nutans	*B.n.nb2*	16.56	24.11	39.96	21.51	21.11	25.30
B. nutans	*B.n.nb3*	48.28	54.86	71.64	20.91	21.13	23.39
B. striata	*B.s.eco1*	19.98	24.82	27.55	17.93	18.40	16.12
B. striata	*B.s.eco2*	22.44	27.46	29.14	21.97	14.97	22.49

B. striata	*B.s.eco2*	21.77	27.60	29.95	10.24	15.03	19.12
Dendrocalamus giganteus	*D.g.nb1*	17.01	26.60	43.80	17.35	19.46	18.88
D. giganteus	*D.g.nb2*	*16.40*	*23.19*	*44.66*	*19.51*	*20.34*	*19.54*
D. giganteus	*D.g.nb3*	18.63	29.64	46.53	15.29	16.66	20.13
D. strictus	*D.s.mdh1*	14.77	16.31	26.82	21.56	21.76	23.19
D. strictus	*D.s.mdh2*	23.66	34.38	35.83	21.46	23.35	25.31
D. strictus	*D.s.mdh3*	14.45	15.67	25.08	17.25	19.99	18.13

4.6. Statistical analysis combining anatomical and biochemical data

Correlation analysis was done between each of the four anatomical parameters and the biochemical estimations of cellulose and lignin levels in order to evaluate the usefulness of the four anatomical parameters to predict superior species/genotypes having suitable ligno-cellulose contents. The biochemical level of cellulose and VB density had a Pearson correlation of 0.505, while the biochemical level of cellulose and metaxylem lumen diameter had a Pearson correlation of -0.531, and the biochemical estimate of cellulose vs ID of FG had a Pearson correlation of 0.776 **(Table 7)**. In a similar manner, the Pearson correlation values between the biochemical level of lignin and the VB density, lumen diameter of metaxylem, and ID for PG were 0.481, -0.398, and 0.507, respectively. For lignin, a modest correlation (0.507) was obtained between the biochemical estimates and the Integrated Density for PG. On the other hand, in case of cellulose, the correlation value between biochemical estimate and ID for FG was higher (0.776). This could be due to a very non-linear relationship between lignin level and ID for PG and consequently, Pearson correlation might not be the best way to represent the association. Therefore, the Spearman's rho was used as a more reliable and non-parametric option for additional investigation. The Spearman's rho revealed a much stronger association between the biochemical level of lignin and the ID for PG (0.519) than between the Vb density and metaxylem lumen diameter (0.415 and 0.389, **Table 7**). All the correlation data found from the analysis were highly significant. By combining data on biochemical estimates of cellulose and lignin and anatomical ID for FG and PG, a summary statistic was calculated, and each species was then given a summary score. Higher value would suggest a high amount of cellulose and lignin for both cases. According to the index score, *B. nutans* (0.483) and *B. bambos* (0.367) had high cellulose levels **(Table 8)**, whereas *B. striata* (-0.60) and *B. balcooa* (-0.30) had low lignin levels **(Table 8)**. Correlation analysis was also performed using all biochemical data acquired from the 63 samples (7 species, 3 genotypes, and 3 internode positions) in order to assess the variation between

biochemically estimated cellulose and lignin contents in bamboo. The obtained Pearson correlation value of 0.45 **(Fig. 5)** suggested strong correlation between these two at high significant level (p-value = $3.117e^{-4}$).

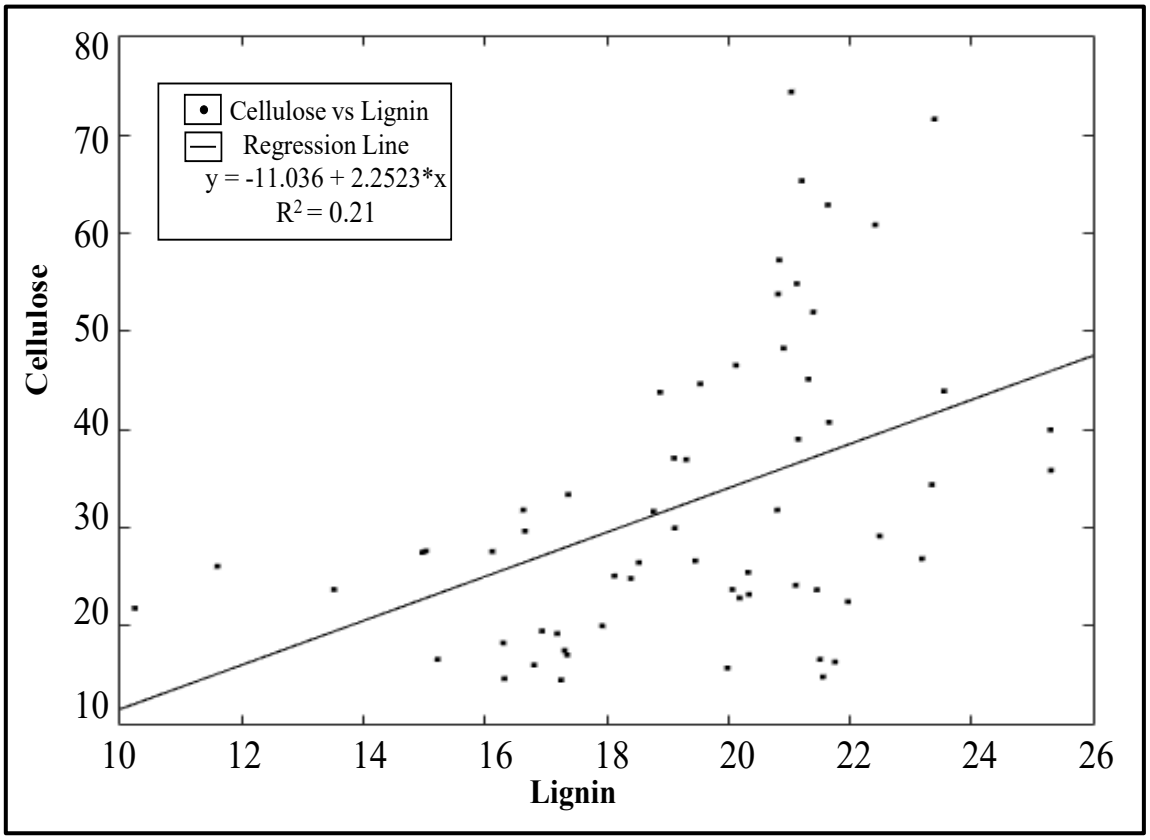

Fig. 5. Correlation analysis between the biochemically estimated cellulose and acid insoluble lignin in seven species of bamboo. For each species, three different genotypes were evaluated.

Table 7. Correlation analysis between biochemical level of cellulose and lignin and selected anatomical parameters and stain intensities.

Biochemical parameter	Correlation (r) with anatomical parameters		
	No. of vascular bundle	Diameter of lumen	Stain intensity
α-cellulose	Pearson correlation (r) = 0.505 (95% confidence interval, 0.290, 0.671) [p-value = $3.272e^{-05}$]	Pearson correlation (r) = -0.530 (95% confidence interval, -0.690, -0.321) [p-value = $1.085e^{-05}$]	Pearson correlation (r) = 0.776 (95% confidence interval, 0.652, 0.860) [p-value = $1.944e^{-13}$]
α-cellulose	Spearman's rho= 0.583 (p-value = $8.086e^{-07}$)	Spearman's rho= -0.602 (p-value = $2.715e^{-07}$)	Spearman's rho= 0.786 (p-value = $6.123e^{-14}$)

Acid insoluble lignin	Pearson correlation (r) = 0.481 (95% confidence interval, 0.261, 0.654) [p-value = $8.502e^{-05}$]	Pearson correlation (r) = -0.397 (95% confidence interval, -0.590, -0.162) [p-value = $1.504e^{-03}$]	Pearson correlation (r) = 0.507 (95% confidence interval, 0.293, 0.673) [p-value = $2.972e^{-05}$]
Acid insoluble lignin	Spearman's rho= 0.415 (p-value = $8.68e^{-04}$)	Spearman's rho= -0.389 (p-value = $2.058e^{-03}$)	Spearman's rho= 0.519 (p-value = $2.393e^{-05}$)

Table 8. Summary statistics to assign a summary score to each species to predict bamboo species that are superior with respect to high cellulose and low lignin contents.

Bamboo species	**Species score with respect to cellulose**	**Species score with respect to lignin**
Bambusa balcooa	-0.133	**-0.300**
B. tulda	-0.033	-0.050
B. bambos	**0.366**	0.166
B. nutans	**0.483**	0.483
B. striata	0.233	0.216
Dendrocalamus giganteus	-0.466	0.100
D. strictus	-0.183	**-0.600**

Objective 3: Identification, phylogenetic comparisons and expression characterization of selected *CesA* homologs in bamboo

4.7. *In silico* identification of *CesA* family genes from all the sequenced genomes of monocotyledonous plants

Coding sequences of *A. thaliana CesA* genes were obtained from TAIR database. Most of these genes have been identified and functionally characterized in Arabidopsis, and therefore were used as queries to BLSATP. At least one homolog for each *AtCesA* could be identified from other related monocot species such as *O. sativa*, *Phyllostachys heterocycla*, *Brachypodium distachyon*, *Zea mays*, *Setaria italica*, *Setaria viridis*, *Sorghum bicolor*, *Hordeum vulgare*, *Panicum virgatum*, *Panicum halii*.

4.8. Selection of candidate *CesA* genes involved in the secondary cell wall formation

Functional characterization of CesA genes has been well established in Arabidopsis. Three CesA genes (*AtCesA1*, *AtCesA3*, and *AtCesA6*) are involved in the development of the

primary cell wall, whereas other three CesA genes (*AtCesA4*, *AtCesA7*, and *AtCesA8*) are linked to the development of the secondary cell wall. Hence, homologous genes encoding *AtCesA4*, *AtCesA7* and *AtCesA8* were identified from *O. sativa.* Similarly, *OsCesA4* (homologous to *AtCesA8*), *OsCesA7* (homologous to *AtCesA4*) and *OsCesA9* (homologous to *AtCesA7*) were identified having role in secondary cell wall development. Additional homologous sequences were identified from the five bamboo species (*P. heterocycla*, *Bonia amplexicaulis*, *Guadua angustifolia*, *Olyra latifolia, Raddia guianensis*) by accessing the Bamboo genome database (Gendbe). For this purpose, *O. sativa* genes were used as the querry sequence. At least one homolog was identified for each species **(Table 9)**.

Table 9. *In silico* identification of homologous *CesA* genes from bamboos. BLASTP analysis was performed where *O. sativa* sequences were used as query. For any individual gene if more than one BLAST hits were obtained, the top BLAST hit was selected for further analyses.

***OsCesA* genes** / **Bamboo species**	***OsCesA4* (Os01g54620)**	***OsCesA7* (Os10g32980)**	***OsCesA9* (Os09g25490)**
Phyllostachys edulis	*PH01000018G0380* *PH01000746G0570*	*PH01000693G0390*	*PH01000924G0590*
Raddia guinensis	*Rgu021448.1*	*Rgu023114.1*	*Rgu023255.1*
Bonia amplexicaulis	*Bam028804.1* *Bam003109.1*	*Bam016116.1* *Bam011661.2*	*Bam036200.1* *Bam027690.2*
Guadua angustifolia	*Gan003698.1* *Gan038367.1*	*Gan027374.1* *Gan013500.1*	*Gan024116.1*
Olyra latifolia	*Ola033738.1*	*Ola003157.1*	*Ola009992.1*

4.9. Identification and molecular characterization of three SCW *CesA* gene homolog from *B. tulda*

In this study, three *CesA* genes had been isolated from *B. tulda* by designing degenerate primers and were sequenced (**Table 10)**. The full length obtained sequences were submitted to NCBI database (*BtCesA4* - **OP121187**). BLASTP analysis indicated that a very high level of homology for *BtCesA4* was obtained in case of the homologous *B. eminensis* (95.43%), *O. brachyantha* (94.84%) and *O. sativa* (94.74%) genes. The *BtCesA4* genomic sequence length is 5227 kb, coding sequence is 2943 nucleotides long and predicted protein is 980 amino acid

long. The predicted protein contains a highly conserved Plant conserved region (PCR) domain, and D, D, D, QxxRW motif.

Table 10. Summary of the sequenced genes obtained from *B. tulda* and their sequence homology analysis. Only top 5 candidates from the blast results were included here.

Obtained genes from *B. tulda*		**Blast Results**			
Genes	**Total Gene Length (bp)**	**Species**	**Identity (%)**	**Query Cover (%)**	**E value**
BtCesA4	5227 (Full length)	*Bambusa eminensis*	95.43	100	0.0
		Oryza brachyantha	94.84	100	0.0
		O. sativa	94.74	100	0.0
		Setaria italica	93.37	99	0.0
		Zea mays	93.22	98	0.0
BtCesA7	2379 (Partial)	*Bambusa eminensis*	98.90	99	0.0
		O. brachyantha	98.08	99	0.0
		O. sativa	97.80	99	0.0
		O. glaberrima	97.80	99	0.0
		Panicum virgatum	97.80	99	0.0
BtCesA9	3576 (Partial)	*Hordeum vulgare*	92.02	99	0.0
		S. italica	91.99	99	0.0
		S. viridis	91.90	99	0.0
		B. distachyon	91.46	100	0.0
		P. hallii	91.02	99	0.0

4.10. Study on the phylogenetic relationship of *B. tulda CesA* gene in comparison to the homologous sequences from related monocot species

To understand the phylogenetic relationships among the *BtCesA4*, *BtCesA7* and *BtCesA9* gene and their homologous monocotyledonous counterparts, NJ trees were calculated using multiple aligned protein sequences **(Fig. 6).** Six major clade was obtained. Homologs of *CesA1*, *CesA4*, *CesA7* and *CesA9* were obtained in respective clusters and *BtCesA4* was along with homologs *CesA4* genes **(Fig. 6)**.

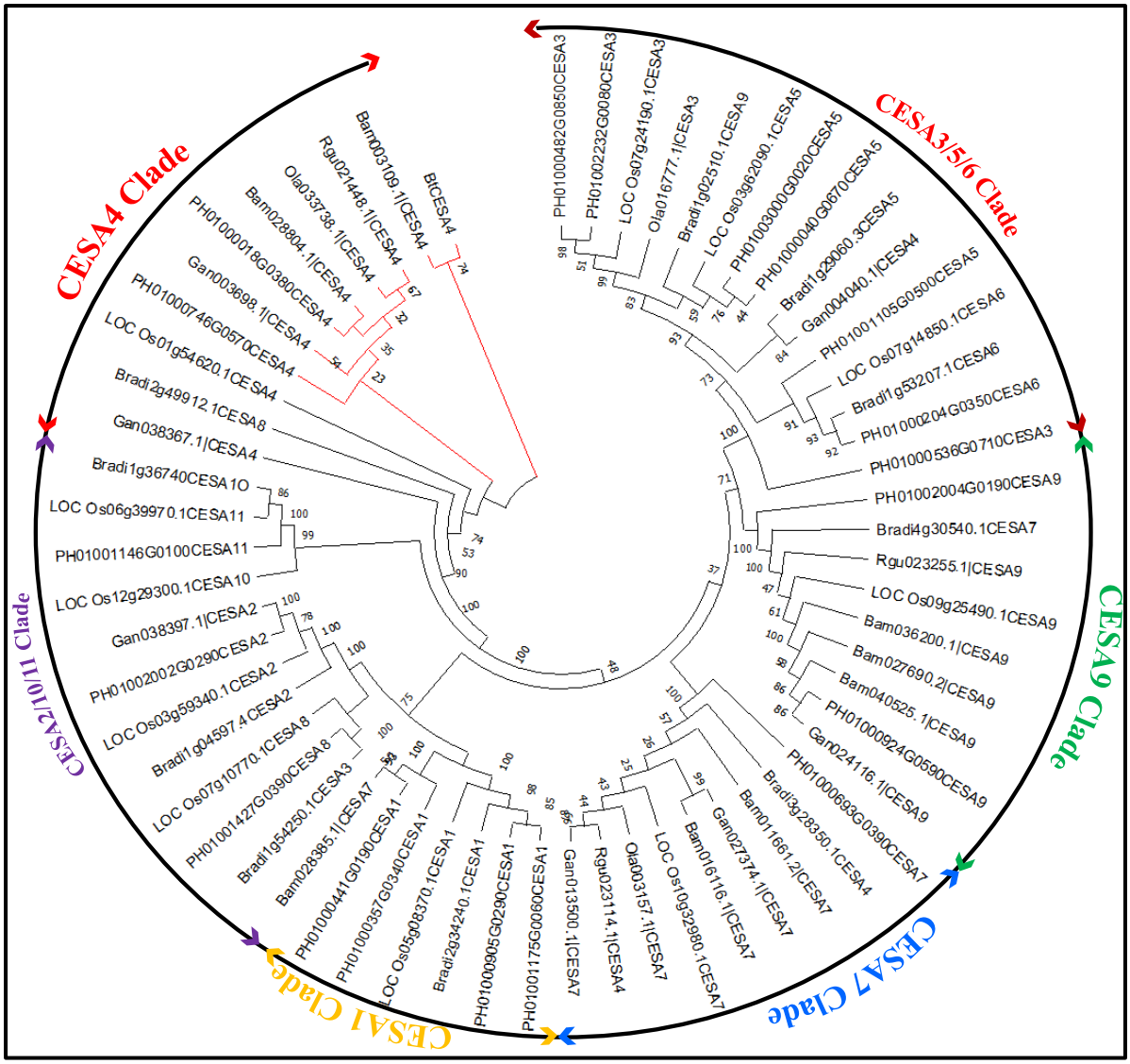

Fig. 6. Phylogenetic analysis of *BtCesA4* gene with other SCW related *CesA* genes from different monocotyledonous Plants. Neighbour Joining (NJ) tree of the ***BtCesA4*** homologs was constructed after multiple alignment of full-length protein sequences in Mega 11.0 and afterward constructing the phylogenetic tree using default parameters and bootstrap value 1000.

4.11. Comparative analyses of predicted gene model structures in relation to those of similar monocotyledonous species

Predicted gene model of *BtCesA4* revealed that the gene featured thirteen exons, and had highest similarity with *OsCesA4* from *O. sativa* **(Fig. 7)**. It was quite similar to *OsCesA4* with respect to the total length, exon number, exon length and intron length. In contrast, the temperate bamboo *P. heterocycla* had lesser number of exons.

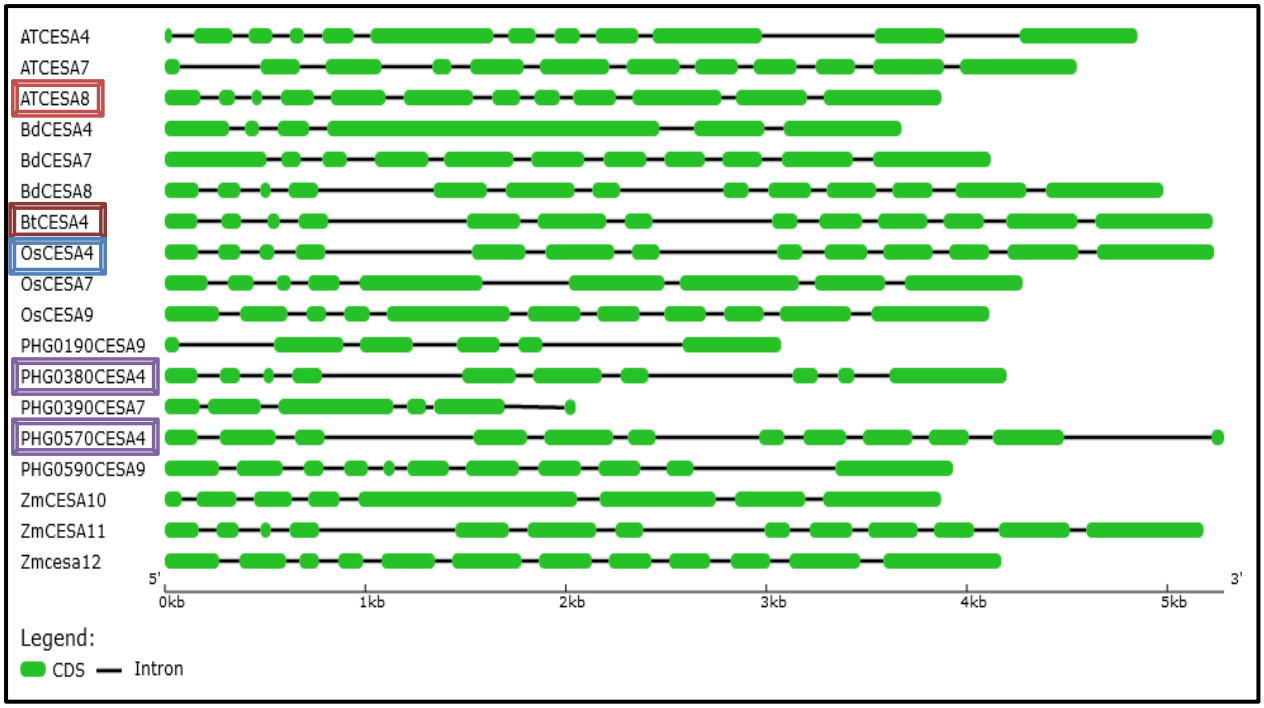

Fig. 7. Gene model structure prediction of *BtCesA4* gene with other SCW related *CesA* genes from different monocotyledonous Plants. *AtCesA8* is marked in red box, *BtCesA4* is marked in brown box, *OsCesA4* is marked in blue box, *PhCesA4* is marked in purple box.

4.12. *In silico* identification of domains present in CesA protein homologs of closely related grasses

Multiple sequence alignment of the predicted BtCesA4 protein in comparison to rice and Brachypodium homologs revealed that the biologically important domains are present and conserved in *B. tulda* also. All the domains starting from N-terminal domain, Zinc finger domain, Variable region 1 and 2, Transmembrane domain 1 and 2, Conserved region 1 and 2, C-terminal domain were present and the amino acid positions of the domains were shown in **Fig. 8**. Amino acid sequences were highly conserved across all the CesA4 protein sequences.

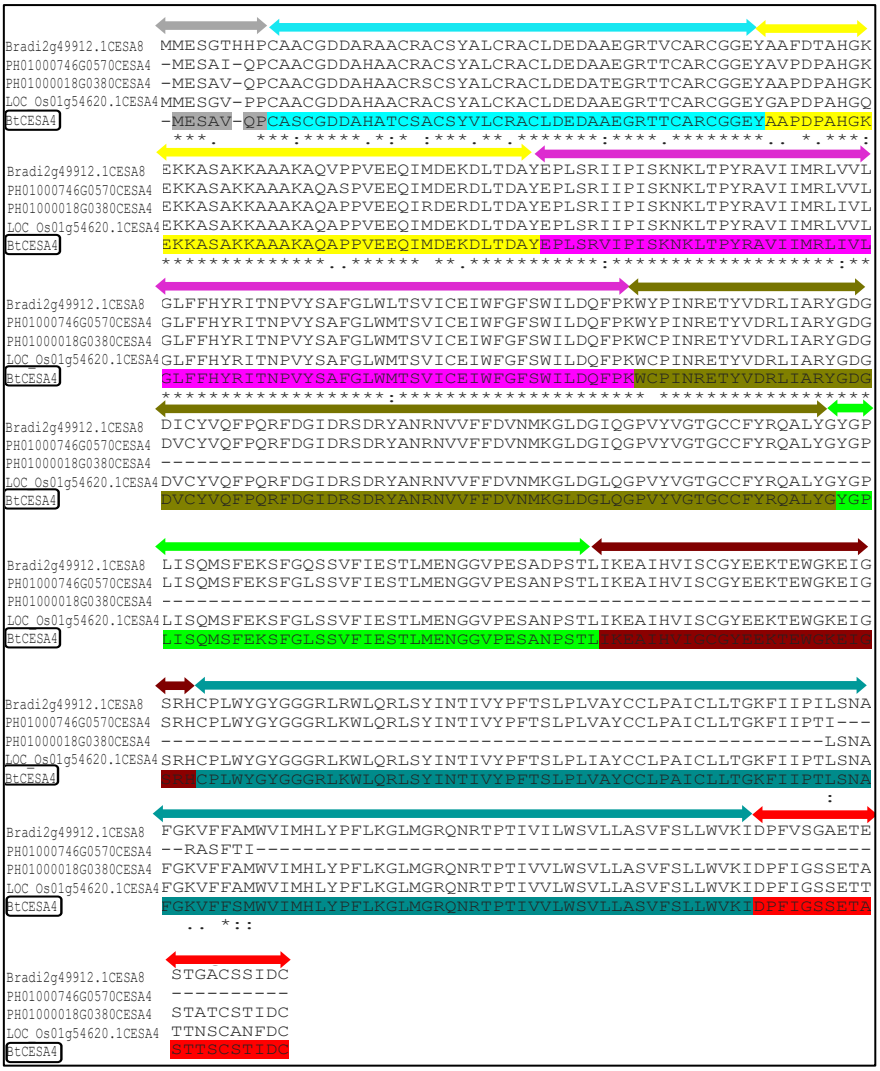

Fig. 8. *In silico* identification of all the domains with the amino acid positions present in BtCesA4 proteins and comparison with phylogenetically related grass species. Homolog identified from *B. tulda* was marked with black box. Abbreviations used: NT- N-terminal domain, ZN – Zinc finger domain, VR1- Variable region 1, TM1- Transmembrane domain 1, CR1 – Conserved region 1, VR2 – Variable region 2, CR2 – Conserved Region 2. TM2 – Transmembrane domain 2, CT- C terminal domain. Bradi – *Brachypodium distachyon*, PH – *Phyllostachys heterocycla*, Os – *Oryza sativa*, Bt – *Bambusa tulda*.

5. Discussion

5.1. Bamboo germplasms having superior ligno-cellulose property may be identified by using a simple, affordable method

Multiple methods have been applied to estimate the amounts of cellulose, hemicellulose, and lignin in various bamboo species that include near infrared spectroscopy (Li et al., 2015), FTIR

macro and micro-spectroscopy (Li et al., 2018), HPLC (Huang et al., 2020), FTIR and NMR (Shi et al., 2019). While each of these methods are sensitive and powerful, there remains the need for a cheaper, more straightforward, and easy-to-perform method. Assessing the integrated density of cellulose or lignin specific dye can be effective to screen a large number of bamboo species/genotype with respect to their ligno-cellulose property. Our analysis had found strong correlation between the biochemical level of cellulose and lignin and the stain integrated density. In addition, lumen diameter of metaxylem was also found to be negatively corelated to the biochemical level of cellulose and lignin. These anatomical parameters can be implemented to assess a large number of bamboo genotypes.

5.2. Cellulose and lignin contents in bamboos are impacted by one another and dependent upon its physiological age

Using data obtained from the 7 species, 3 genotypes, and 3 internode sites (Total = 63 data), correlation analysis was done to determine whether any relationships exist between the cellulose and lignin contents in bamboo. Despite the biological variability present across the samples, the resulting Pearson correlation was reasonably high (0.45, **Fig. 5**). It implies that cellulose content impacts the amount of lignin in bamboo and *vice versa*. Similar observation has been made in other plants including switchgrass (Xu et al., 2020) and sugarcane (Chandel et al., 2009). Many environmental and biological variations such as altitude, temperature, water and light availability, and plant's physiological age affect the biosynthesis of cellulose and lignin (Llerena et al., 2019). The current study comprised three different internode tissues, such as basal, middle, and apical internodes, in order to assess the impact of physiological age on cellulose-lignin levels. The amount of cellulose and lignin levels were more in the developed tissues (e.g., basal internode) than the less developed one (e.g., apical internode). This finding was consistent with other plants such as Pennisetum (Zhang et al., 2020), miscanthus (da Costa et al., 2014).

5.3. *Bambusa bambos* could be used as the model bamboo plant for the LCB based biofuel industry

Identification of a "superior species" based on the biomass accumulation, carbon content, and cellulose vs. lignin profile was one major objective of this study. However, only three species (*P. edulis*, *B. chungii*, *B. cacharensis*) have so far been studied to understand biomass accumulation in bamboos (Huang et al., 2020; Jin et al., 2021). Among them, the temperate bamboo *Phyllostachys edulis* had primarily been targeted to understand cellulose-lignin biology (Li et al., 2015; Li et al., 2018; Ling et al., 2021). In spite of a vast majority of bamboo vegetation available in tropical regions, no studies have been conducted on any tropical species. In this study, *B. bambos* was predicted as the superior species due to having highest

cellulose and relatively lower amount of lignin. Additional anatomical, biochemical and analytical data such as the vascular anatomy, lowest thermal stability (320.67 °C), crystallinity index (41.05%) and residual char content also supported *B. bambos* as the superior species and ideal raw material for the biofuel industry (Biswas et al., 2022). In order to obtain biofuel from LCB, pretreatment is necessary to extract cellulose from the cellulose-hemicellulose-lignin matrix. However, it is costly and often an energy demanding procedure. The LCBs obtained from *B. bambos* is expected to require pretreatment at a lesser extent and hence is more favourable from the energy and environmental perspective.

5.4. Genetic engineering of *CesA* genes might lead to enhanced biomass accumulation or lower crystallinity index

To enhance plant growth, biomass output, and wood quality, constant attempts are being made to modify the cellulose synthesis process *in planta* by targeting the *CesA* genes (Maleki et al., 2020). Altering *CesA* gene functions by employing genetic modification tools such as CRISPR Cas9 or RNA interference may increase the availability of cellulose per unit area of biomasses by lowering the cellulose crystallinity and thereby exposing lignin to effective pre-treatment processes (Singh et al., 2022). Although, the first draft genome sequencing of a bamboo species has been conducted quite sometimes ago (Peng et al., 2013), only a few studies have been conducted to study the sequences and expression of genes related to cellulose biosynthesis (Chen et al. 2010). So, a major aim of this study was to sequence *CesA* genes that are related to major biomass accumulation. In that respect, Arabidopsis *CesA4, 7* and *8* are important since they are involved in secondary cell wall formation. Therefore, in this study, *B. tulda* homolog of *At.CesA* 4, 7 and 8 have been identified.

6. Significance and Future direction

This study investigated seven bamboo species that are commonly growing in West Bengal, India based on their anatomical, biochemical and physicochemical properties to assess their potential for lignocellulosic biomass production. The investigation identified *B. bambos* as superior species for future biofuel research. However, future research should concentrate on pretreatment and saccharification reactions in order to utilize identified bamboo LCBs for bioenergy. Additionally, it is important to gather sequence information of cellulose biosynthesis genes for the purpose of future genetic engineering. Therefore, three *CesA* genes related to secondary cell wall formation had also been characterized from *B. tulda*. These genes can be targeted in future to create transgenic plants by promoting cellulose biosynthesis or by modifying regulation in plant cellulose biosynthetic pathway.

7. References:

Agbor VB, Cicek N, Sparling R, Berlin A, Levin DB (2011) Biomass pretreatment: fundamentals toward application. Biotechnol Adv 29 (6):675-685. doi:10.1016/j.biotechadv.2011.05.005

Akinyemi AB, Omoniyi ET, Onuzulike G (2020) Effect of microwave assisted alkali pretreatment and other pretreatment methods on some properties of bamboo fibre reinforced cement composites. Construction and Building Materials 245:118405. doi:https://doi.org/10.1016/j.conbuildmat.2020.118405

Arioli T, Peng L, Betzner AS, Burn J, Wittke W, Herth W, Camilleri C, Hofte H, Plazinski J, Birch R, Cork A, Glover J, Redmond J, Williamson RE (1998) Molecular analysis of cellulose biosynthesis in Arabidopsis. Science 279 (5351):717-720. doi:10.1126/science.279.5351.717

Barakat A, de Vries H, Rouau X (2013) Dry fractionation process as an important step in current and future lignocellulose biorefineries: a review. Bioresour Technol 134:362-373. doi:10.1016/j.biortech.2013.01.169

Basak M, Dutta S, Biswas S, Chakraborty S, Sarkar A, Rahaman T, Dey S, Biswas P, Das M (2021) Genomic insights into growth and development of bamboos: what have we learnt and what more to discover? Trees 35 (6):1771-1791. doi:10.1007/s00468-021-02197-6

Bhattacharya S, Ghosh JS, Sahoo DK, Dey N, Pal AJAoFS (2010) Screening of superior fiber-quality-traits among wild accessions of Bambusa balcooa: efficient and non-invasive evaluation of fiber developmental stages. 67 (6):611-611. doi:10.1051/forest/2010024

Biswal D, Shinkhede S, Mandavgane SA (2021) Bamboo Valorization by Fermentation and Enzyme Treatment. In: Thatoi H, Mohapatra S, Das SK (eds) Bioprospecting of Enzymes in Industry, Healthcare and Sustainable Environment. Springer Singapore, Singapore, pp 87-101. doi:10.1007/978-981-33-4195-1_4

Biswas S, Rahaman T, Gupta P, Mitra R, Dutta S, Kharlyngdoh E, Guha S, Ganguly J, Pal A, Das M (2022) Cellulose and lignin profiling in seven, economically important bamboo species of India by anatomical, biochemical, FTIR spectroscopy and thermogravimetric analysis. Biomass and Bioenergy 158:106362. doi:https://doi.org/10.1016/j.biombioe.2022.106362

Chakraborty S, Dutta S, Biswas P, Das M (2019) Identification of candidate reference genes in tropical bamboos stable across species, tissues, and developmental stages. Biol Plant 63 (1):253-261. doi:10.32615/bp.2019.029

Chandel AK, Narasu ML, Chandrasekhar G, Manikyam A, Rao LV (2009) Use of *Saccharum spontaneum* (wild sugarcane) as biomaterial for cell immobilization and modulated ethanol production by thermotolerant *Saccharomyces cerevisiae* VS3. Bioresour Technol 100 (8):2404-2410. doi:10.1016/j.biortech.2008.11.014

Chen C-Y, Hsieh M-H, Yang C-C, Lin C-S, Wang A-Y (2010) Analysis of the cellulose synthase genes associated with primary cell wall synthesis in Bambusa oldhamii. Phytochemistry 71 (11):1270-1279. doi:https://doi.org/10.1016/j.phytochem.2010.05.011

Chin KL, Ibrahim S, Hakeem KR, San H'ng P, Lee SH, Lila MAM (2017) Bioenergy production from bamboo: potential source from Malaysia's perspective. BioResources 12 (3):6844-6867

da Costa, R.M., Lee, S.J., Allison, G.G., Hazen, S.P., Winters, A. and Bosch, M., 2014. Genotype, development and tissue-derived variation of cell-wall properties in the lignocellulosic energy crop Miscanthus. *Annals of Botany*, *114*(6), pp.1265-1277.

Das M, Bhattacharya S, Pal A (2005) Generation and characterization of SCARs by cloning and sequencing of RAPD products: a strategy for species-specific marker development in bamboo. Ann Bot 95 (5):835-841. doi:10.1093/aob/mci088

Das M, Bhattacharya S, Singh P, Filgueiras TS, Pal A (2008) Bamboo Taxonomy and Diversity in the Era of Molecular Markers. In: Advances in Botanical Research, vol 47. Academic Press, pp 225-268. doi:https://doi.org/10.1016/S0065-2296(08)00005-0

Delmer DP, Haigler CH (2002) The Regulation of Metabolic Flux to Cellulose, a Major Sink for Carbon in Plants. Metabolic Engineering 4 (1):22-28. doi:https://doi.org/10.1006/mben.2001.0206
He C, Cui K, Zhang J, Duan A, Zeng Y (2013) Next-generation sequencing-based mRNA and microRNA expression profiling analysis revealed pathways involved in the rapid growth of developing culms in Moso bamboo. BMC Plant Biology 13 (1):119. doi:10.1186/1471-2229-13-119
Hosseinaei O, Wang S, Enayati AA, Rials TG (2012) Effects of hemicellulose extraction on properties of wood flour and wood–plastic composites. Composites Part A: Applied Science and Manufacturing 43 (4):686-694. doi:https://doi.org/10.1016/j.compositesa.2012.01.007
Huang L, Yang Z, Li M, Liu Z, Qin C, Nie S, Yao S (2020) Effect of Pre-Corrected pH on the Carbohydrate Hydrolysis of Bamboo during Hydrothermal Pretreatment. Polymers 12 (3):612
Isikgor FH, Becer CR (2015) Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. Polym Chem 6 (25):4497-4559. doi:10.1039/C5PY00263J
Jin Y, Liu J, Yang H, Shi Z, Zhao P, Yang J (2021) Improving enzymatic saccharification and ethanol production of bamboo residues with sulfomethylation-aided phosphoric acid pretreatment. Ind Crop Prod 170:113733. doi:https://doi.org/10.1016/j.indcrop.2021.113733
Kang Q, Appels L, Tan T, Dewil R (2014) Bioethanol from lignocellulosic biomass: current findings determine research priorities. The Scientific World Journal 2014:298153. doi:10.1155/2014/298153
Kimura S, Laosinchai W, Itoh T, Cui X, Linder CR, Brown RM, Jr. (1999) Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant vigna angularis. The Plant cell 11 (11):2075-2086
Kubovský I, Kačíková D, Kačík F (2020) Structural Changes of Oak Wood Main Components Caused by Thermal Modification. Polymers 12 (2):485
Li X, Sun C, Zhou B, He Y (2015) Determination of Hemicellulose, Cellulose and Lignin in Moso Bamboo by Near Infrared Spectroscopy. Sci Rep 5 (1):17210. doi:10.1038/srep17210
Li X, Wei Y, Xu J, Xu N, He Y (2018) Quantitative visualization of lignocellulose components in transverse sections of moso bamboo based on FTIR macro- and micro-spectroscopy coupled with chemometrics. Biotechnol Biofuels 11 (1):263. doi:10.1186/s13068-018-1251-4
Lin J, He X, Hu Y, Kuang T, Ceulemans R (2002) Lignification and lignin heterogeneity for various age classes of bamboo (*Phyllostachys pubescens*) stems. Physiol Plant 114 (2):296-302. doi:10.1034/j.1399-3054.2002.1140216.x
Ling Z, Tang W, Su Y, Shao L, Wang P, Ren Y, Huang C, Lai C, Yong Q (2021) Promoting enzymatic hydrolysis of aggregated bamboo crystalline cellulose by fast microwave-assisted dicarboxylic acid deep eutectic solvents pretreatments. Bioresour Technol 333:125122. doi:10.1016/j.biortech.2021.125122
Llerena JPP, Figueiredo R, Brito MdS, Kiyota E, Mayer JLS, Araujo P, Schimpl FC, Dama M, Pauly M, Mazzafera P (2019) Deposition of lignin in four species of Saccharum. Sci Rep 9 (1):5877. doi:10.1038/s41598-019-42350-3
Maleki SS, Mohammadi K, Movahedi A, Wu F, Ji KS (2020) Increase in Cell Wall Thickening and Biomass Production by Overexpression of PmCesA2 in Poplar. Front Plant Sci 11. doi:10.3389/fpls.2020.00110
Mølhøj M, Pagant S, Höfte H (2002) Towards Understanding the Role of Membrane-bound Endo-β-1,4-glucanases in Cellulose Biosynthesis. Plant and Cell Physiology 43 (12):1399-1406. doi:10.1093/pcp/pcf163
Morais S, Morag E, Barak Y, Goldman D, Hadar Y, Lamed R, Shoham Y, Wilson DB, Bayer EA (2012) Deconstruction of lignocellulose into soluble sugars by native and designer cellulosomes. MBio 3 (6). doi:10.1128/mBio.00508-12

Peng Z, Lu Y, Li L, Zhao Q, Feng Q, Gao Z, Lu H, Hu T, Yao N, Liu K, Li Y, Fan D, Guo Y, Li W, Lu Y, Weng Q, Zhou C, Zhang L, Huang T, Zhao Y, Zhu C, Liu X, Yang X, Wang T, Miao K, Zhuang C, Cao X, Tang W, Liu G, Liu Y, Chen J, Liu Z, Yuan L, Liu Z, Huang X, Lu T, Fei B, Ning Z, Han B, Jiang Z (2013) The draft genome of the fast-growing non-timber forest species moso bamboo (Phyllostachys heterocycla). Nat Genet 45 (4):456-461. doi:10.1038/ng.2569

Pradhan Mitra P, Loqué D (2014) Histochemical staining of *Arabidopsis thaliana* secondary cell wall elements. J Vis Exp (87). doi:10.3791/51381

Read SM, Bacic T (2002) Prime Time for Cellulose. 295 (5552):59-60. doi:10.1126/science.1068155 %J Science

Richmond TJGB (2000) Higher plant cellulose synthases. 1 (4):reviews3001.3001. doi:10.1186/gb-2000-1-4-reviews3001

Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET, Eliceiri KW (2017) ImageJ2: ImageJ for the next generation of scientific image data. BMC Bioinformatics 18 (1):529-529. doi:10.1186/s12859-017-1934-z

Ruel K, Nishiyama Y, Joseleau J-P (2012) Crystalline and amorphous cellulose in the secondary walls of Arabidopsis. Plant Science 193-194:48-61. doi:https://doi.org/10.1016/j.plantsci.2012.05.008

Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative CT method. Nature Protocols 3 (6):1101-1108. doi:10.1038/nprot.2008.73

Segal L, Creely JJ, Martin AE, Conrad CM (1959) An Empirical Method for Estimating the Degree of Crystallinity of Native Cellulose Using the X-Ray Diffractometer. Textile Research Journal 29 (10):786-794. doi:10.1177/004051755902901003

Sharma R, Wahono J, Baral H (2018) Bamboo as an Alternative Bioenergy Crop and Powerful Ally for Land Restoration in Indonesia. Sustainability 10 (12):4367

Shi Z, Xu G, Deng J, Dong M, Murugadoss V, Liu C, Shao Q, Wu S, Guo Z (2019) Structural characterization of lignin from D. sinicus by FTIR and NMR techniques. Green Chemistry Letters and Reviews 12 (3):235-243. doi:10.1080/17518253.2019.1627428

Singh P, Dimri R, Chamoli P, Jha SK (2022) ADVANCED PRE-PROCESSING STRATEGIES FOR LIGNOCELLULOSIC BIOMASS APPLYING GENETIC ENGINEERING AND NANOTECHNOLOGY TO INCREASE BIOETHANOL YIELD. Journal of Pharmaceutical Negative Results:883-890

Updegraff DM (1969) Semimicro determination of cellulose inbiological materials. Anal Biochem 32 (3):420-424. doi:https://doi.org/10.1016/S0003-2697(69)80009-6

Vazquez-Cooz I, Meyer RW (2002) A differential staining method to identify lignified and unlignified tissues. Biotech Histochem 77 (5-6):277-282. doi:10.1080/bih.77.5-6.277.282

Xu K, Shi Z, Lyu J, Zhang Q, Zhong T, Du G, Wang S (2020) Effects of hydrothermal pretreatment on nano-mechanical property of switchgrass cell wall and on energy consumption of isolated lignin-coated cellulose nanofibrils by mechanical grinding. Ind Crop Prod 149:112317. doi:https://doi.org/10.1016/j.indcrop.2020.112317

Yousuf A, Pirozzi D, Sannino F (2020) Chapter 1 - Fundamentals of lignocellulosic biomass. In: Yousuf A, Pirozzi D, Sannino F (eds) Lignocellulosic Biomass to Liquid Biofuels. Academic Press, pp 1-15. doi:https://doi.org/10.1016/B978-0-12-815936-1.00001-0

Zabed H, Sahu JN, Boyce AN, Faruq G (2016) Fuel ethanol production from lignocellulosic biomass: An overview on feedstocks and technological approaches. Renewable and Sustainable Energy Reviews 66:751-774. doi:https://doi.org/10.1016/j.rser.2016.08.038

Zhang W, Zhang S, Lu X, Li C, Liu X, Dong G, Xia T (2020) Tissue-specific Transcriptome analysis reveals lignocellulose synthesis regulation in elephant grass (*Pennisetum purpureum* Schum). BMC Plant Biology 20 (1):528. doi:10.1186/s12870-020-02735-3

Zhao X, Zhang L, Liu D (2012) Biomass recalcitrance. Part I: the chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose. 6 (4):465-482. doi:doi:10.1002/bbb.1331

Zheng X, Lin S, Fu H, Wan Y, Ding Y (2020) The Bamboo Flowering Cycle Sheds Light on Flowering Diversity. Front Plant Sci 11 (381). doi:10.3389/fpls.2020.00381